Audio Engineering Society

Convention Paper

Presented at the AES 161st Convention
2026 October 30-November 1, Nashville, TN, USA

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Fast Time-Varying Exponentiated Convolution Methods for Generative Direction Dependent Reverberation

Yuancheng Luo[1]

[1]*NuSpace Audio, Cambridge, MA, USA*

Correspondence should be addressed to Yuancheng Luo (luoyuancheng@gmail.com)

ABSTRACT

Spherical harmonic encoded acoustic sound-fields capture directional characteristics of room impulse responses that are useful for accurate spatial audio reproduction. However, high costs of multi-microphone measurements and numerical simulations motivate alternative data-set augmentation and synthetic data generation methods that supplement small collections. This paper introduces time-varying exponentiated convolution methods that transform both Gaussian noise and impulse responses into reverberation and modified spectral-decay fields respectively. We derive two recursive and fast convolution algorithms that extend into the spherical harmonic domain, model smooth reverberation time distributions with non-stationary Gaussian processes, and realize an optimal filter design. Experiments evaluate computational performance, and validate out-of-distribution generated impulse responses.

1 Introduction

Spatial room impulse responses (SRIRs) that decompose acoustic fields over bounded spaces have found numerous applications in spatial audio reproduction over both physical and virtual loudspeaker arrays for auralization. In optimal experimental design, *hardware auralization* of microphone and loudspeaker topology enables performance evaluations on various downstream processors such as beam-forming, acoustic tuning, speech enhancement, and sound-source localization. Spherical harmonic (SH) [1, 2, 3] formatted SRIRs such as Ambisonics [4], are well-suited for auralization as their rotational properties and separability from the hardware's geometry increase the modeling capacity of diverse experimental conditions; SRIRs can be integrated with freely rotated plane-wave decompositions of a transducer's free-field acoustic responses. However, acquiring large and varied SRIR data-sets remains a challenge due to practical costs of measurement and numerical simulation. We therefore investigate synthetic and data augmentation methods of SH-SRIRs as to complement data collection efforts:

Approximate RIR models such as SH domain image-source method [5, 6, 7] are efficient alternatives to expensive numerical simulations, capable of generating large quantity of SH-RIRs by varying simplified shoe-box room geometries and wall-material absorption properties. Time-varying smoothing kernels [8, 9] and low-pass filters [10, 11] can augment simulated RIRs with spectral-decay characteristics of distance-based atmosphere attenuation. Directional filtering using SH filter-banks [12, 13] can modify the directivity of sound-fields. Our work synthesizes these approaches by generalizing time-varying filtering for exponentiating finite-impulse-response (FIR) kernels that control for direction dependent spectral-decay under SH expansions. As a result, we can both augment and generate

SH-SRIRs with direction dependent spectral-decay distributions sampled from a Gaussian process (GP) [14].

Our paper is organized as follows: Section 2 introduces two time-varying exponentiating convolution methods for modifying IR spectral-decay without adding delay. Section 3 extends our methods into the SH domain for directional filtering and reverberation; we present a filter optimization program and a non-stationary GP model of reverberation time distributions. Section 4 evaluates several experiments with reverberation time sampling, filter fitting sensitivity, and SH-SRIR generation. Section 5 concludes the work. Our implementation is available online[1].

2 Exponentiating Convolutions

Let us consider the following generalization of the convolution function $f(\boldsymbol{h}, \boldsymbol{g})$ between a fixed FIR $\boldsymbol{h}$ and a time-varying FIR filter $\boldsymbol{g}(m)$ undergoing exponential-convolution at each time index m:

$$\begin{aligned} \boldsymbol{h} = [h_1, \dots, h_M] \in \mathbb{R}^M, \quad \boldsymbol{g} = [g_1, \dots, g_N] \in \mathbb{R}^N, \\ \boldsymbol{g}(m) = \boldsymbol{g}(m-1) \circledast \boldsymbol{g} = \circledast_{k=1}^{m} \boldsymbol{g} \in \mathbb{R}^{m(N-1)+1}, \end{aligned} \tag{1}$$

where $\circledast$ is the convolution operator. The n^{th} sample of the time-varying convolution function $f(\boldsymbol{h}, \boldsymbol{g})$ follows

$$f[n] = \sum_{m=1}^{MN} h[n-m+1]\, g(n-m+1)[m], \tag{2}$$

where its non-zero array elements are indexed via

$$\begin{aligned} h[m] &= \begin{cases} h_m, & 1 \le m \le M \\ 0, & \text{Otherwise} \end{cases}, \quad \text{Fixed FIR} \\ g[n] &= \begin{cases} g_n, & 1 \le n \le N \\ 0, & \text{Otherwise} \end{cases}, \quad \text{Exp. Filter} \\ g(m)[n] &= \begin{cases} \sum_{k=1}^{mN} g[n-k+1]\, g(m-1)[k], & m > 1 \\ g[n], & m = 1 \\ 0, & m < 1 \end{cases}. \end{aligned} \tag{3}$$

The later samples in $\boldsymbol{h}$ are convolved with samples of higher-order filters $\boldsymbol{g}(m)$ under exponentiation. Thus, direct computation of $f(\boldsymbol{h}, \boldsymbol{g})$ can be expressed in terms of summations over sample-weighted exponentiating convolutions in Algorithm 1. Its asymptotic run-time costs $\mathrm{O}\left(M^2 N\right)$ however are prohibitive as the exponentiating filter $\boldsymbol{g}(m)$ grows by $N-1$ samples per time-step, yielding a size of $M(N-1)+1$ when directly convolving with filter $\boldsymbol{h}$ in the time-domain.

[1] https://github.com/yluo1/SH-IRT

Algorithm 1 Direct Exponentiating Convolution

procedure DIREXPCONV($\boldsymbol{h}$, $\boldsymbol{g}$)
Input: $\boldsymbol{h} \in \mathbb{R}^M$ Fixed FIR
Input: $\boldsymbol{g} \in \mathbb{R}^N$ Exponentiating FIR
Output: $\boldsymbol{f} \in \mathbb{R}^{MN}$ Accumulating FIR
  $\boldsymbol{f} \leftarrow \boldsymbol{0}^{MN}$ ▷ Initialize accumulator
  $\boldsymbol{u} \leftarrow \boldsymbol{g}$ ▷ Initialize Exp. FIR
  for $m = 1 \dots M$ **do** ▷ Sum convolutions over $\boldsymbol{h}$
    $f[m \text{ to } mN] \leftarrow f[m \text{ to } mN] + h[m]\,\boldsymbol{u}$
    $\boldsymbol{u} \leftarrow \boldsymbol{u} \circledast \boldsymbol{g}$ ▷ Exp. convolution via Eq. (1)
  return $\boldsymbol{f}$

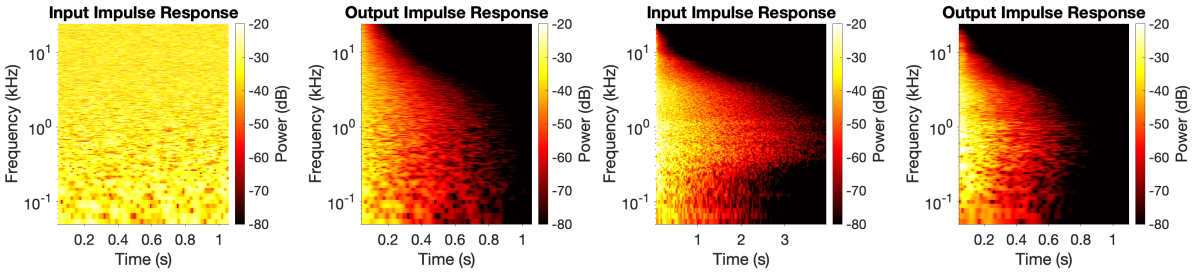
Fortunately, we can obtain a faster algorithm by deriving the transfer functions of both $\boldsymbol{g}(m)$ and $f(\boldsymbol{h}, \boldsymbol{g})$. The transfer function $G(m,z) = G^m(z)$ of the exponentiating filter $\boldsymbol{g}(m)$ over the Z domain is the exponentiated transfer function $G(z)$ of $\boldsymbol{g}$. The transfer function $F_M(z)$ of $f(\boldsymbol{h}, \boldsymbol{g})$ following Algorithm 1 can be expressed by

$$\begin{aligned} F_M(z) &= \sum_{m=0}^{M-1} h_{m+1} z^{-m} G^{m+1}(z), \quad \text{Time-vary Conv.} \\ G^m(z) &= \left(\sum_{n=0}^{N-1} g_{n+1} z^{-n} \right)^m, \quad \text{Exp. Filter} \end{aligned} \tag{4}$$

where we observe that the exponentiated filter's transfer function $G^m(z) = G^{\lfloor m/2 \rfloor}(z) G^{m-\lfloor m/2 \rfloor}(z)$ can be recursively divided into halves of m via the floor function $\lfloor * \rfloor$. Moreover, the transfer function $F_M(z)$ can be recursively divided into upper and lower sample halves of $\boldsymbol{h}$, and expressed as the following recurrence relation:

$$\begin{aligned} F_M(z, \boldsymbol{h}) &= \sum_{m=1}^{M} h[m] z^{-(m-1)} G^m(z) \\ &= \begin{cases} \underline{\underline{F_{\underline{M}}}}(z) + \overline{F_{\underline{M}}}(z)\, G^{\underline{M}}(z)\, z^{-\underline{M}}, & M > 1 \\ h[m] G(z), & M = 1 \end{cases}, \\ \overline{F_{\underline{M}}}(z) &= F_{M-\underline{M}}\left(z, h\left[(\underline{M}+1), \dots, M\right]\right), \\ \underline{\underline{F_{\underline{M}}}}(z) &= F_{\underline{M}}\left(z, h\left[1, \dots, \underline{M}\right]\right), \quad \underline{M} = \left\lfloor \frac{M}{2} \right\rfloor, \end{aligned} \tag{5}$$

where $F_M(z, \boldsymbol{h})$ decomposes into summations of two transfer functions over $\boldsymbol{h}$ separated w.r.t. pivot index $\underline{M}$. The lower-half transfer function $\underline{\underline{F_{\underline{M}}}}(z)$ remains unchanged, whereas the upper-half $\overline{F_{\underline{M}}}(z)$ is multiplied by the exponentiated filter's transfer function $G^{\underline{M}}(z)$ and delay $z^{-\underline{M}}$. The latter constitutes the anticipated work required to recursively compute the transfer function when sub-divided upto a maximum depth of $\log_2(M)$.



(a) Input Gaussian noise $\boldsymbol{h}$ to output $f(\boldsymbol{h},\boldsymbol{g})$ spectrograms

(b) Transform spectral-decay of input RIR $\boldsymbol{h}$ to output $f(\boldsymbol{h},\boldsymbol{g})$

Fig. 1: Exponentiating filtering convolves the fixed input FIR $\boldsymbol{h}$ with the time-varying filter $\boldsymbol{g}$ undergoing sample-wise exponential-convolution of Eqs. (1), (2) to achieve target frequency dependent attenuation rates.

Algorithm 2 Optimized Exponentiating Convolution

procedure OPTEXPCONV($\boldsymbol{h}$, $\boldsymbol{g}$)
Input: $\boldsymbol{h} \in \mathbb{R}^M$ Fixed FIR
Input: $\boldsymbol{g} \in \mathbb{R}^N$ Exponentiating FIR
Output: $\boldsymbol{f} \in \mathbb{R}^{MN}$ Accumulating FIR
  $\underline{M} \leftarrow \lfloor \frac{M}{2} \rfloor$
  if $M \leq M_{min}$ **then** ▷ M_{min} is nominally small
    $\boldsymbol{f} \leftarrow$ DirExpConv($\boldsymbol{h}$, $\boldsymbol{g}$) ▷ via Algo. 1
  else ▷ Compute two halves via Eq. (5)
    $\overline{\boldsymbol{f}} \leftarrow$ OptExpConv($h[\underline{M}+1$ to $M]$, $\boldsymbol{g}$)
    $\underline{\boldsymbol{f}} \leftarrow$ OptExpConv($h[1$ to $\underline{M}]$, $\boldsymbol{g}$)
    $\boldsymbol{f} \leftarrow \left[\underline{\boldsymbol{f}}, \boldsymbol{0}^{(M-\underline{M})N}\right] + \left[\boldsymbol{0}^{\underline{M}}, \overline{\boldsymbol{f}} \circledast \boldsymbol{g}(\underline{M})\right]$
  return $\boldsymbol{f}$

We implement the recurrence relation of Eq. (5) in our optimized exponentiating convolution Algorithm 2. The computational work in each recursive call is the convolution operation between $\overline{\boldsymbol{f}}$ and $\boldsymbol{g}(\underline{M})$, which can be efficiently performed in the frequency domain via the Fast Fourier transform (FFT) at $\mathrm{O}(MN\log(MN))$ computation costs; nominally small partition sizes revert to the direct method in Algorithm 1 when its costs are cheaper than the FFT implementation. Furthermore, the problem sub-divides into equal-sized partitions which simplifies the complexity analysis; integrating the work over all partitions and depths gives the total asymptotic computation cost of $\mathrm{O}\left(MN\left((\log(M))^2+\log(M)\log(N)\right)\right)$. Note that in practice, $N \lll M$ as we design filters $\boldsymbol{g}$ to be near minimum-phase with magnitude responses that produce expected reverberation time profiles. The examples in Fig. 1 use a $N=4$ order filter $\boldsymbol{g}$ to shape frequency dependent attenuation rates of an $M=72000$ length Gaussian noise $\boldsymbol{h}$ in Fig. 1a, and $M=192000$ length RIR in Fig. 1b. Next, we present an alternative but more expensive auto-regressive formulation of the exponentiating convolution process.

Recursive Auto-regressive Convolution: Consider the modified exponentiating convolutions of Eq. (2), by which $f[n]$ are samples of a first-order auto-regressive function $f_m[n]$ where $f[n]=f_M[n]$. The auto-regressive function's recurrence relations are given by

$$f_m[n] = \begin{cases} f_{m-1}[n], & m>1, n<m \\ \sum_{k=m}^{n} f_{m-1}[k]\, g[n-k+1], & m>1, n\geq m \\ \sum_{k=1}^{M} h[k]\, g[n-k+1], & m=1 \\ h[n], & m=0 \end{cases}, \tag{6}$$

where the initial state of $f_0[n]$ is FIR $\boldsymbol{h}$, and current samples of $f_m[n]$ are convolutions of partial samples $f_{m-1}[n]$ from the previous time-step with filter $\boldsymbol{g}$. We give the direct implementation in Algorithm 3 where a sliding window of $f[n]$ recursively convolves with $\boldsymbol{g}$.

Algorithm 3 Direct Recursive Convolution

procedure DIRRECURCONV(m, $\boldsymbol{h}$, $\boldsymbol{g}$)
Input: m Trim index
Input: $\boldsymbol{h} \in \mathbb{R}^M$ Fixed FIR
Input: $\boldsymbol{g} \in \mathbb{R}^N$ Exponentiating FIR
Output: $\boldsymbol{f} \in \mathbb{R}^{M+m(N-1)}$ Accumulating FIR
  $\boldsymbol{f} \leftarrow \boldsymbol{0}^{M+m(N-1)}$ ▷ Initialize accumulator
  $f[1$ to $M] \leftarrow \boldsymbol{h}$ ▷ Initialize head of $\boldsymbol{f}$ to $\boldsymbol{h}$
  for $i=1\ldots m$ **do** ▷ Conv. $\boldsymbol{g}$ over trimmed $\boldsymbol{f}$
    $\boldsymbol{u} \leftarrow f[i$ to $M+(N-1)(i-1)] \circledast \boldsymbol{g}$
    $f[i$ to $M+(N-1)i] \leftarrow \boldsymbol{u}$
  return $\boldsymbol{f}$

The asymptotic run-time costs $\mathrm{O}\left(M^2N\right)$ of the direct Algorithm 3 are as prohibitive as the direct exponenti-

Algorithm 4 Optimized Recursive Convolution

procedure OPTRECURCONV(m, $\boldsymbol{h}$, $\boldsymbol{g}$)
Input: m Trim index
Input: $\boldsymbol{h} \in \mathbb{R}^M$ Fixed FIR
Input: $\boldsymbol{g} \in \mathbb{R}^N$ Exponentiating FIR
Output: $\boldsymbol{f} \in \mathbb{R}^{M+\underline{m}(N-1)}$ Accumulating FIR
  $\underline{m} \leftarrow \lfloor \frac{m}{2} \rfloor$
  $\overline{m} \leftarrow M + \underline{m}(N-1)$
  if $m \leq M_{min}$ **then** ▷ M_{min} is nominally small
    $\boldsymbol{f} \leftarrow$ DirRecurConv(m, $\boldsymbol{h}$, $\boldsymbol{g}$) ▷ via Algo. 3
  else
    $\overline{\boldsymbol{f}} \leftarrow h[\underline{m}+1 \text{ to } M] \circledast \boldsymbol{g}(\underline{m})$
    $\underline{\boldsymbol{f}} \leftarrow$ OptRecurConv($h[1 \text{ to } \underline{m}]$, $\boldsymbol{g}$)
    $\boldsymbol{u} \leftarrow [\underline{\boldsymbol{f}}, \boldsymbol{0}^{M-\underline{m}}] + [\boldsymbol{0}^{\underline{m}}, \overline{\boldsymbol{f}}]$
    $\boldsymbol{v} \leftarrow$ OptRecurConv($m-\underline{m}, u[\underline{m}+1 \text{ to } \overline{m}]$)
    $\boldsymbol{f} \leftarrow [u[1 \text{ to } \underline{m}], \boldsymbol{v}]$
  return $\boldsymbol{f}$

ating convolution in Algorithm 1, and produce similar outputs for minimum-phase $\boldsymbol{g}$. Fortunately, we derive an analogous but more expensive partitioned formulation to Algorithm 2 in Algorithm 4. The crucial difference lies in its two partition sizes, $\underline{m}$ and $M+\underline{m}(N-2)$, which do not remain proportional across sub-problems as the latter depends on the sizes of both inputs $\boldsymbol{h}$ and $\boldsymbol{g}$. The work per recursive call is the convolution operation, with computational costs $\mathrm{O}(MN\log(MN))$ via FFT; approximate max partition depth is $\log_2 M$. The total asymptotic costs of the largest partitioned trace is $\mathrm{O}\left(MN^2\left((\log(M))^2+\log(M)\log(N)\right)\right)$.

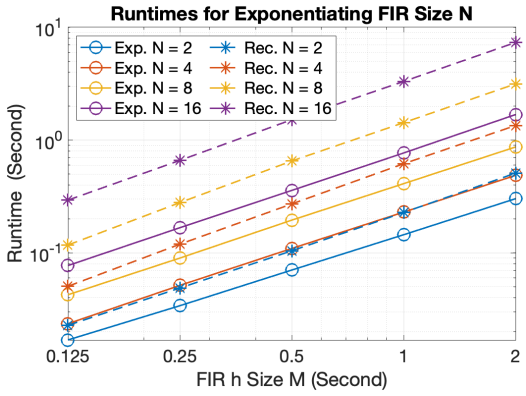


Fig. 2: Runtimes (Matlab, Mac M1, 20 run averages) for exponentiating and recursive Algorithms 2, 4 of varying FIR sizes at 48 kHz sampling rate

Lastly, we compare runtime performances of the optimized exponentiating and recursive auto-regressive convolution methods for increasing FIR sizes of $\boldsymbol{h}$, and $\boldsymbol{g}$ in Fig. 2. The time-invariant FIR $\boldsymbol{h}$ is sampled from a normal distribution, and the exponentiating filter $\boldsymbol{g}$ is fitted to a randomized spectral-decay target. Both methods exhibit log-linear runtimes w.r.t. size M of the input FIR $\boldsymbol{h}$, but only the exponentiating convolution has log-linear runtimes w.r.t. size N of input filters $\boldsymbol{g}$.

3 Directional Extensions

We now consider extending time-varying exponentiating convolution $f(\boldsymbol{h},\boldsymbol{g})$ to SH expansions of multi-mic pressure fields. In the general case of multichannel IRs, let us apply $f(\boldsymbol{h}_i,\boldsymbol{g})$ to the i^{th} channel $\boldsymbol{h}_i^T$ transposed along the i^{th} row of the multichannel row-matrix $\boldsymbol{H}$:

$$\boldsymbol{H} = [\boldsymbol{h}_1,\ldots,\boldsymbol{h}_{N_H}]^T \in \mathbb{R}^{N_H\times M}, \quad \boldsymbol{h}_i \in \mathbb{R}^{M\times 1}. \tag{7}$$

Note that multichannel matrices can be linearly transformed under change of bases such as the real or complex SHs, and stored in encoded formats (e.g. Ambisonics B-format). Therefore, it is useful to show that exponentiating convolutions $f(\boldsymbol{h},\boldsymbol{g})$ commute under all linear transformations $A:\mathbb{C}^{N_H}\rightarrow\mathbb{C}^{N_A}$ given by

$$A\left(\begin{bmatrix} f(\boldsymbol{h}_1,\boldsymbol{g}) \\ \vdots \\ f(\boldsymbol{h}_{N_H},\boldsymbol{g}) \end{bmatrix}\right) = \begin{bmatrix} f(\tilde{\boldsymbol{h}}_1,\boldsymbol{g}) \\ \vdots \\ f(\tilde{\boldsymbol{h}}_{N_H},\boldsymbol{g}) \end{bmatrix} \in \mathbb{C}^{N_A\times MN}, \tag{8}$$

where $A(\boldsymbol{H}) = \left[\tilde{\boldsymbol{h}}_1,\ldots,\tilde{\boldsymbol{h}}_{N_H}\right]^T$ is the linear transformation of the multichannel matrix $\boldsymbol{H}$. The proof and derivation is given in Appendix Eqs. (22), (23), which follow from the observation that A linearly transforms only the time-invariant term $\boldsymbol{h}$ in both commutations. Thus, we can directly call Algorithms 2, 4 on any pressure field encoded channels $\tilde{\boldsymbol{h}}$, and recover their time-varying convolved pressure responses by decoding the channels under any linear transformation. Let us now define directional filtering in terms of a linear transformation of channels encoded along SH bases.

Directional Filtering: Suppose an instantaneous pressure field function $p_C(r,\theta,\phi) \approx \boldsymbol{Y}(\theta,\phi)\boldsymbol{C}(r)$ over time samples r and spherical coordinates (θ,ϕ), and a time-invariant directional attenuation function $D(\theta,\phi) \approx \boldsymbol{Y}(\theta,\phi)\boldsymbol{D}$ are expanded over finite-order SH bases:

$$\boldsymbol{Y}(\theta,\phi) = \left[Y_0^0, Y_1^{-1}, Y_1^0, Y_1^1, \ldots, Y_{L_*}^{L_*}\right] \in \mathbb{C}^{1\times(L_*+1)^2},$$
$$Y_l^m(\theta,\phi) = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}\,P_l^m(\cos\theta)e^{im\phi}, \tag{9}$$

where θ, ϕ are co-latitude and azimuth respectively, and the associated Legendre polynomials $P_l^m(\cos\theta)$ are indexed via expansion order m bounded by $-l \le m \le l$, and degree l bounded by $0 \le l \le L_*$ for max-degree L_*. The expansion weights, given by the column matrix $\boldsymbol{C} = [\boldsymbol{C}(1), \ldots, \boldsymbol{C}(M)] \in \mathbb{C}^{N_C \times M}$ for $N_C = (L_C+1)^2$, and $\boldsymbol{D} \in \mathbb{C}^{N_D \times 1}$ for $N_D = (L_D+1)^2$, are projections of the pressure and attenuation functions on the SH bases respectively.

Let us define a directional weighted pressure field $p_E(r,\theta,\phi)$ by the following product of attenuation and pressure field functions under SH expansions:

$$p_E(r,\theta,\phi) = D(\theta,\phi)\, p_C(r,\theta,\phi) \approx \boldsymbol{Y}(\theta,\phi)\boldsymbol{E}, \quad (10)$$

where its SH expansion can be expressed in terms of $\boldsymbol{E} \in \mathbb{C}^{N_E \times 1}$ weighted SH bases for $N_E = (L_D + L_C + 1)^2$ as the SH product operator has closure under addition [15, 16]. Furthermore, the SH product operator can be decomposed into a linear transform $A(\boldsymbol{D})$ of the directional filter coefficients $\boldsymbol{D}$ represented by the *transfer matrix* [17] $\boldsymbol{A} \in \mathbb{C}^{N_E \times (L_C+1)^2}$ whose entrants are given in Appendix Eq. (24). Thus, the weighted expansion coefficients $\boldsymbol{E} = \boldsymbol{A}\boldsymbol{C}$ simplify into a matrix-vector product of the pressure field's expansion coefficients.

In practice, the product function and SH weights in Eq. (10) can be efficiently computed via forward and inverse fast spherical harmonic transform (FSHT) [18]:

$$\begin{aligned}\boldsymbol{E} &= \boldsymbol{Y}^{-1}(\boldsymbol{\theta},\boldsymbol{\phi})\left(\boldsymbol{Y}_D(\boldsymbol{\theta},\boldsymbol{\phi})\boldsymbol{D} \odot \boldsymbol{Y}_C(\boldsymbol{\theta},\boldsymbol{\phi})\boldsymbol{C}(r)\right) \\ &= \mathrm{FSHT}\left(\mathrm{IFSHT}(\boldsymbol{D}) \odot \mathrm{IFSHT}(\boldsymbol{C}(r))\right) \\ &\approx \mathrm{FSHT}\left(D(\boldsymbol{\theta},\boldsymbol{\phi}) \odot \mathrm{IFSHT}(\boldsymbol{C}(r))\right), \end{aligned} \quad (11)$$

where $\odot$ is element-wise multiplication, and $\boldsymbol{\theta}, \boldsymbol{\phi} \in \mathbb{R}^{N_E}$ are spherical coordinates of a uniform and high-resolution grid sampled over a sphere. Therefore, the square matrix $\boldsymbol{Y}(\boldsymbol{\theta},\boldsymbol{\phi}) \in \mathbb{C}^{N_E \times N_E}$ of SH evaluations at the grid coordinates is well-conditioned and invertible. The rectangular matrices $\boldsymbol{Y}_D(\boldsymbol{\theta},\boldsymbol{\phi})$, $\boldsymbol{Y}_C(\boldsymbol{\theta},\boldsymbol{\phi})$ truncate $\boldsymbol{Y}(\boldsymbol{\theta},\boldsymbol{\phi})$ to the first N_D, N_C columns respectively. For illustration, we specify a directional attenuation function by expanding a squared exponential kernel function over the spherical coordinates, and whose product with a randomized pressure field is shown in Fig. 3.

We now extend time-varying convolutions to SH bases for the case of direction independent filters $\boldsymbol{g}$. The equivalences in Eq. (11) and linearity of Eq. (8) permits us to compute either $f(\boldsymbol{C},\boldsymbol{g})$ directly on each of the N_C rows of the pressure field's SH expansion coefficients $\boldsymbol{C}$, or alternatively $f(\boldsymbol{P},\boldsymbol{g})$ on each of the N_E rows on the spherical grid coordinates of the pressure field matrix $\boldsymbol{P} = \boldsymbol{Y}_C(\boldsymbol{\theta},\boldsymbol{\phi})\boldsymbol{C} \in \mathbb{R}^{N_E \times M}$. Therefore, the following time-varying convolved SH expansion matrix $\tilde{\boldsymbol{E}} \in \mathbb{C}^{N_E \times MN}$ has two equivalences:

$$\tilde{\boldsymbol{C}} = \begin{bmatrix} \tilde{\boldsymbol{C}}_1 \\ \vdots \\ \tilde{\boldsymbol{C}}_{(L_C+1)^2} \end{bmatrix}, \quad \tilde{\boldsymbol{C}}_i = f([C_{i1}, \ldots, C_{iM}], \boldsymbol{g}),$$

$$\tilde{\boldsymbol{P}} = \begin{bmatrix} \tilde{\boldsymbol{P}}_1 \\ \vdots \\ \tilde{\boldsymbol{P}}_{(L_D+L_C+1)^2} \end{bmatrix}, \quad \tilde{\boldsymbol{P}}_i = f([P_{i1}, \ldots, P_{iM}], \boldsymbol{g}),$$

$$\begin{aligned}\tilde{\boldsymbol{E}} &= \boldsymbol{Y}^{-1}(\boldsymbol{\theta},\boldsymbol{\phi})\left(\boldsymbol{Y}_D(\boldsymbol{\theta},\boldsymbol{\phi})\boldsymbol{D}\mathbf{1}^T \odot \boldsymbol{Y}_C(\boldsymbol{\theta},\boldsymbol{\phi})\tilde{\boldsymbol{C}}\right), \\ &= \boldsymbol{Y}^{-1}(\boldsymbol{\theta},\boldsymbol{\phi})\left(\boldsymbol{Y}_D(\boldsymbol{\theta},\boldsymbol{\phi})\boldsymbol{D}\mathbf{1}^T \odot \tilde{\boldsymbol{P}}\right), \end{aligned} \quad (12)$$

where computing over the SH expansion coefficients $\tilde{\boldsymbol{C}}$ requires $L_D^2 + 2L_D(L_C+1)$ fewer evaluations of f than computing over the sample pressure field matrix $\boldsymbol{P}$. Next, we consider the case of time-varying filtering with direction dependent filters $\boldsymbol{g}(\theta,\phi)$.

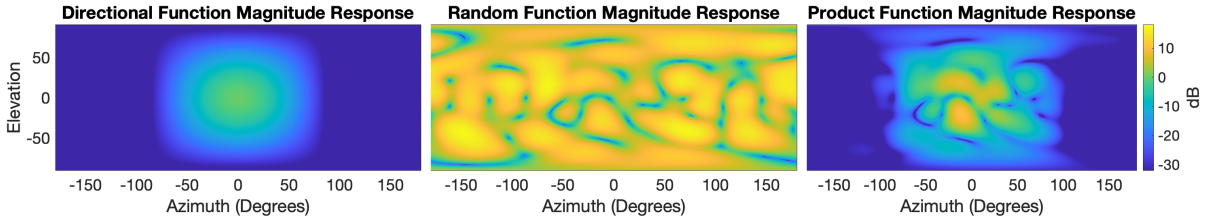
Directional Reverberation: Suppose we allow exponentiating filters $\boldsymbol{g}$ for pressure field $\tilde{\boldsymbol{P}}$ in Eq. (12) to vary as if sampled over spherical coordinates via $\boldsymbol{g}(\theta,\phi)$. The rows of $\tilde{\boldsymbol{P}}_i$ are pressure time-series on the spherical grid coordinates $(\theta_i,\phi_i) \in (\boldsymbol{\theta},\boldsymbol{\phi})$ become

$$\tilde{\boldsymbol{P}}_i = f([P_{i1}, \ldots, P_{iM}], \boldsymbol{g}(\theta_i,\phi_i)), \quad (13)$$

whereby rows of pressure matrix $\boldsymbol{P}$ at the grid coordinates undergo direction dependent convolution. In the case of unit sized $(N=1)$ filters $\boldsymbol{g}(\theta,\phi)$, the time-varying convolution process is equivalent to exponential-decaying directional filtering of Eq. (10), where the direction attenuation function $D(\theta,\phi) = \boldsymbol{g}^m(\theta,\phi)$ at time-sample m exponentiates a single coefficient. For higher-order filters, we require the filter's frequency response $G(\omega,\theta,\phi) = G(z,\theta,\phi)|\, z = e^{j\omega}$ for angular frequency ω, to smoothly vary over spherical coordinates as to recover a low-order expansion matrix $\tilde{\boldsymbol{E}}$. Consider the gradient of the exponentiated filter's magnitude frequency responses $|G^m(\omega,\theta,\phi)|$ w.r.t. spherical coordinates given by

$$\nabla|G^m(\omega,\theta,\phi)| = m|G(\omega,\theta,\phi)|^{m-1}\nabla|G(\omega,\theta,\phi)|. \quad (14)$$

The gradient $\nabla|G(\omega,\theta,\phi)|$ should be differentiable for non-zero magnitudes, and the exponentiating term



(a) Direction attenuation function $D(\theta,\phi)$ **(b)** Pressure field function $p_C(r,\theta,\phi)$ **(c)** Weighted pressure field $p_E(r,\theta,\phi)$

Fig. 3: We expand the kernel function of chordal distances [16] given by $k(d) = e^{-|d|^2/(2\ell)}$ at the expansion center $(\theta_c = 90°, \phi_c = 0°)$ upto max-degree $L_D = 7$ in Fig. 3a, where d is the Euclidean distance between spherical coordinates (θ,ϕ) and (θ_c,ϕ_c) on the unit sphere . The product of the directional function and randomized pressure field in Fig. 3b (max-degree $L_C = 7$) attenuates the directional responses in Fig. 3c.

grows unbounded unless we restrict the filter's magnitude responses $0 < |G(\omega,\theta,\phi)| \leq \tau < 1$. Thus, we model $|G(\omega,\theta,\phi)|$ as follows:

Let the magnitude response $|G(\omega,\theta,\phi)|$ achieve a target reverberation time of $T_{60}(\omega,\theta,\phi)$ seconds for 60 decibel (dB) attenuation. Equating the exponentiated magnitude response $|G^m(\omega,\theta,\phi)| = 10^{-3}$ for the attenuation's sample time $m = F_s T_{60}(\omega,\theta,\phi)$ at the sampling rate F_s yields the following magnitude dB target:

$$|G(\omega,\theta,\phi)|_{dB} = \frac{-60}{F_s\, \mathrm{T}_{60}(\omega,\theta,\phi)}, \tag{15}$$

by which substituting into the exponential term $m|G(\omega,\theta,\phi)|^{m-1}$ in Eq. (14) gives the following upper bound on the exponential magnitude component:

$$\begin{aligned} m_* &= \arg\max_m m\,|G(\omega,\theta,\phi)|^{m-1} = \frac{F_s T_{60}(\omega,\theta,\phi)}{3\log 10}, \\ \Rightarrow m\,|G(\omega,\theta,\phi)|^{m-1} &\leq 10^{\frac{3}{F_s T_{60}(\omega,\theta,\phi)}} \frac{F_s T_{60}(\omega,\theta,\phi)}{(3\log 10)e}, \end{aligned} \tag{16}$$

which is approximately linear w.r.t. $T_{60}(\omega,\theta,\phi)$. Thus, longer reverberation times induce larger gradients in Eq. (14) that must be offset by smaller Euclidean norms of the gradient component $\|\nabla|G(\omega,\theta,\phi)|\|$ such that larger $T_{60}(\omega,\theta,\phi)$ are smoother over spherical coordinates. Note that in practice, we observe longer reverberation times at lower frequencies in RIRs, and model distributions of $T_{60}(\omega,\theta,\phi)$ whose smoothness in the spherical coordinates vary inversely with frequency. We present a modified non-stationary GP [19] for modeling $T_{60}(\omega,\theta,\phi)$ distributions in the next section.

With regards to smooth phase components of the exponentiated filter's transfer function $G^m(\omega,\theta,\phi)$, the group delay of $G(\omega,\theta,\phi)$ should be minimized as the exponentiation operation multiplies the phase response $\angle G^m(\omega,\theta,\phi) = m\angle G(\omega,\theta,\phi)$. We therefore specify the filter's response to be near minimum-phase when subjected to constraints on its upper-bound magnitude $|G(\omega,\theta,\phi)| \leq \tau$, which ensures attenuation across all frequencies and decreases the gradient in Eq. (14). The minimum-phase response following the real-cepstrum method [20, 21] can be derived from the desired magnitude responses of Eq. (15) at $\overline{N}$ number of uniformly spaced frequencies between DC and F_s. We enforce the latter magnitude bounds via quadratic constrained least-squares fitting of the N-tap FIR filter $\boldsymbol{g}$ to $\overline{N}$ number of minimum-phase frequency response targets.

Consider the following second-order cone program:

$$\begin{aligned} &(\lambda_*, \boldsymbol{g}_*) = \arg\min_{(\lambda,\boldsymbol{g})} \lambda, \quad \text{s.t.} \quad \|\boldsymbol{W}\boldsymbol{g} - \boldsymbol{b}\| \leq \lambda, \quad \lambda \geq 0, \\ &\left\|\boldsymbol{w}^T(k)\,\boldsymbol{g}\right\| \leq \tau, \quad \forall k \in \left\{0,\ldots,\left\lfloor \frac{K}{2} \right\rfloor\right\}, \quad \text{Constraints} \\ &w_n(k) = e^{-j2\pi k(n-1)/K}, \quad W_{kn} = e^{-j2\pi(k-1)(n-1)/\overline{N}}, \end{aligned} \tag{17}$$

where $\boldsymbol{W} \in \mathbb{C}^{\overline{N}\times N}$ is an $\overline{N}$-sized discrete Fourier transform matrix truncated to the first N columns, and $\boldsymbol{b} = [G(z_1),\ldots,G(z_{\overline{N}})]^T \in \mathbb{C}^{\overline{N}\times 1}$ the vector of target minimum-phase responses at frequencies $\boldsymbol{\omega} = \{z_1,\ldots,z_{\overline{N}}\}$ where $z_n = e^{-j2\pi(n-1)/\overline{N}}$. Minimizing the non-negative variable λ in the objective function minimizes the error norm between the filter and target responses; we can truncate $\boldsymbol{W}$, $\boldsymbol{b}$ to the first $\lfloor \overline{N}/2 \rfloor + 1$ rows such that the errors span only the frequencies between DC and Nyquist. The magnitude responses of $\boldsymbol{g}$ are constrained over K uniform and densely spaced angular frequencies evaluated at the Fourier bases

$\boldsymbol{w}^T(k) = [w_1(k), \ldots, w_N(k)] \in \mathbb{C}^{1\times N}$. Therefore, we can fit fewer coefficients to the number of target responses and constraints as $N \leq \overline{N} \ll K$.

Gaussian Process Model: Let us now model a distribution of smooth $T_{60}(\omega,\theta,\phi)$ functions sampled from a GP with model supports given by the following observed logarithmic reverberation time vector $\boldsymbol{y} = [y_1, \ldots, y_S]^T \in \mathbb{R}^{S\times 1}$ at the set of frequency and spherical coordinates $\boldsymbol{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_S\}$, $\boldsymbol{x}_n = \left(\underline{\omega_n}, \underline{\theta_n}, \underline{\phi_n}\right)$:

$$y_n = \log T_{60}(\underline{\omega_n}, \underline{\theta_n}, \underline{\phi_n}) - \mu(\underline{\omega_n}), \tag{18}$$

where $\boldsymbol{y}$ is centered over a prior log-mean power-law reverberation time function $\mu(\omega) = \log \overline{T}_{60}(\omega)$ of ordinary frequency $\nu(\omega)$ given by

$$\overline{T}_{60}(\omega) = \alpha\,(\nu(\omega))^{-\beta}, \quad \nu(\omega) = \frac{\angle\omega}{2\pi}, \tag{19}$$

such that the scale α and power β parameters are non-negative. The mean function detrends the log reverberation time over frequency, and the earlier log-transformation ensures that sampled reverberation time functions from the following GP distributions are positive: We can sample $T_{60}(\omega,\theta,\phi)$ functions over the input set $\boldsymbol{X}_*$ of size S_* from either the GP prior normal distribution $\mathcal{N}(\boldsymbol{\mu}(\boldsymbol{X}_*), \boldsymbol{k}(\boldsymbol{X}_*,\boldsymbol{X}_*))$ or the predictive normal distribution $\mathcal{N}(\boldsymbol{\mu}_*(\boldsymbol{X}_*), \boldsymbol{\Sigma}_*(\boldsymbol{X}_*))$ conditioned on observed $\boldsymbol{X}$, $\boldsymbol{y}$. The means and covariances follow

$$\begin{aligned}
\boldsymbol{\mu}(\boldsymbol{X}_*) &= [\mu(\omega_1), \ldots, \mu(\omega_{S_*})]^T, && \text{Prior}\\
\boldsymbol{\mu}_*(\boldsymbol{X}_*) &= \boldsymbol{\mu}(\boldsymbol{X}_*) + \boldsymbol{k}(\boldsymbol{X}_*,\boldsymbol{X})\,\underline{\boldsymbol{K}}^{-1}\boldsymbol{y}, && \text{Posterior}\\
\boldsymbol{\Sigma}_*(\boldsymbol{X}_*) &= \boldsymbol{k}(\boldsymbol{X}_*,\boldsymbol{X}_*) - \boldsymbol{k}(\boldsymbol{X}_*,\boldsymbol{X})\,\underline{\boldsymbol{K}}^{-1}\boldsymbol{k}(\boldsymbol{X},\boldsymbol{X}_*),\\
\underline{\boldsymbol{K}} &= \boldsymbol{k}(\boldsymbol{X},\boldsymbol{X}) + \boldsymbol{\Lambda} \in \mathbb{R}^{S\times S},
\end{aligned} \tag{20}$$

where entrants of $\boldsymbol{k}(\boldsymbol{X},\boldsymbol{X}_*) \in \mathbb{R}^{S\times S_*}$ are covariance functions $k_{ij} = k(\boldsymbol{x}_i,\boldsymbol{x}_j)$ of tuples in $\boldsymbol{X}$, $\boldsymbol{X}_*$, and $\boldsymbol{\Lambda}$ is the diagonal matrix of noise variances.

The covariance function $k(\boldsymbol{x},\boldsymbol{x}')$ determines the smoothness of functions drawn from a GP w.r.t. similarity between the inputs $\boldsymbol{x}$ and $\boldsymbol{x}'$. We consider the following covariance function that is stationary in the spherical coordinates (θ,ϕ), and non-stationary in wavelengths $\lambda = \ell/(\nu(\omega))^{\gamma}$ where ℓ, γ, σ are velocity, power, and scaling hyper-parameters respectively:

$$k(\boldsymbol{x},\boldsymbol{x}') = \sigma^2\left(\frac{2\lambda\lambda'}{\lambda^2+\lambda'^2}\right)^{\frac{3}{2}}\exp\left(-\frac{d^2(\theta,\phi,\theta',\phi')}{(\lambda^2+\lambda'^2)/2}\right), \tag{21}$$

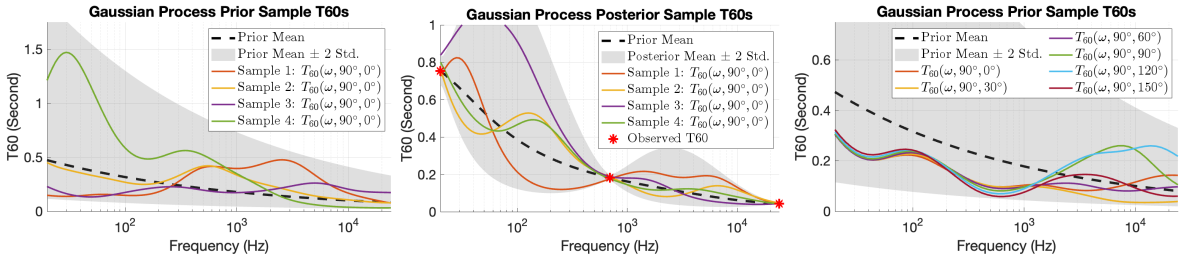
$$d^2(\theta,\phi,\theta',\phi') = 2\sin\left(\frac{\left|\cos^{-1}\left(\boldsymbol{v}^T\boldsymbol{v}'\right)\right|}{2}\right) = \left\|\boldsymbol{v}-\boldsymbol{v}'\right\|_2,$$

where $\boldsymbol{v},\boldsymbol{v}' \in \mathbb{R}^{3\times 1}$ the Cartesian unit-directions of the spherical coordinates (θ,ϕ), (θ',ϕ') respectively, and d is the chordal distance between the points on a unit-sphere. Note that $k(\boldsymbol{x},\boldsymbol{x}')$ is a valid non-stationary squared exponential covariance function for the 3-dimensional case in [19]. The non-stationarity in wavelengths induces a slower decaying exponential over the chordal distance for larger λ. As a result, reverberation times sampled at longer wavelengths in lower frequencies are smoother over the spherical coordinates.

4 Experiments

Gaussian Process T60 Sampling: Let the T_{60} sampling grid $\{\boldsymbol{\omega}\times(\boldsymbol{\theta},\boldsymbol{\phi})\}$ be the Cartesian product between the set of frequencies for fitting $\boldsymbol{g}$ in Eq. (17), and the uniform spherical coordinate grid for recovering SH expansion coefficients $\tilde{\boldsymbol{E}}$ in Eq. (12). We specify the set of sampling inputs $\boldsymbol{X}_* = \left\{\boldsymbol{x}_{11}, \ldots, \boldsymbol{x}_{\overline{N}N_E}\right\}$, where $\boldsymbol{x}_{mn} = (\omega_m,\theta_n,\phi_n)$, and draw sample $T_{60}(\omega,\theta,\phi)$ functions from the GP prior and predictive distribution in Eq. (20) as shown in Fig. 4. The power-law function's parameters $\alpha = 1$, $\beta = 0.25$ yield a tapering GP prior mean $\mu(\omega)$ and $\overline{T}_{60}$ in Eq. (19), and narrows the variance at higher frequencies due to the log-transformation $\log T_{60}$ in Eq. (18) and covariance function in Eq. (21) ($\sigma = \sqrt{2}/2$, $\gamma = 2/3$, $\ell = 343$). Conditioning the GP on several observed T_{60} values at one spherical coordinate yields a predictive distribution with variances that tighten near observed $\boldsymbol{\omega}$, and revert to the prior when afar. Jointly sampling several T_{60} functions across different spherical coordinates exhibits positive co-varying times at lower frequencies due to larger mean-squared wavelengths λ in the non-stationary covariance of Eq. (21).

Exponentiating Filter Sensitivity: Under exponentiation, we expect the transfer functions $G^m(z)$ of $\boldsymbol{g}$ in Eq. (4) to amplify the filter's fitting errors from Eq. (17) as magnitude dB and group-delay responses are multiplied under exponentiation. We show that reducing the filter size N far below the number of target T_{60} responses over frequency introduces large errors. For illustration, consider the following $T_{60} = [4, 0.25, 4, 4, 0.5]$ second targets at uniform spaced frequencies between DC and Nyquist at 48 kHz sampling rate in Fig. 5. Specifying the target log-magnitude responses following Eq. (15) and reflecting over Nyquist yields $N = 8$ tap minimum-phase targets, but which its unconstrained filter responses exceed unity between



(a) Sample T_{60} functions drawn from the GP prior at a common spherical coordinate

(b) Sample T_{60} functions drawn from the GP posterior conditioned on 4 observations

(c) Evaluation grid's T_{60} functions drawn from the GP prior on the horizontal plane

Fig. 4: Sample T_{60} functions that are independently drawn from the GP prior in Fig. 4a vary over a wide range in time/frequency, and narrows when drawn from the GP posterior predictive distribution in Fig. 4b. The T_{60} functions jointly sampled over various spherical coordinates on the same grid in Fig. 4c remain correlated in lower frequencies, and accords with the design of our non-stationary covariance function in frequency.

$12-16$ kHz. The constrained filters' fitted responses for $N \in \{8,6\}$ satisfy the unity bounds $\tau = 1$ over the frequency range and have low fitting errors. Further decreasing N results in both significant deviations from intended target responses, and larger variations across fitted responses over the spherical coordinates. We can decrease the error variance by increasing the T_{60} resolution in frequency over the GP evaluation grid.

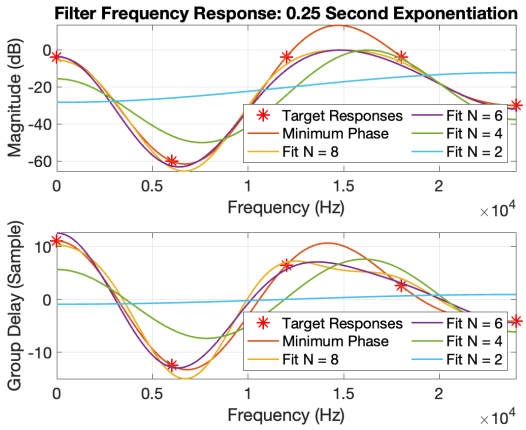
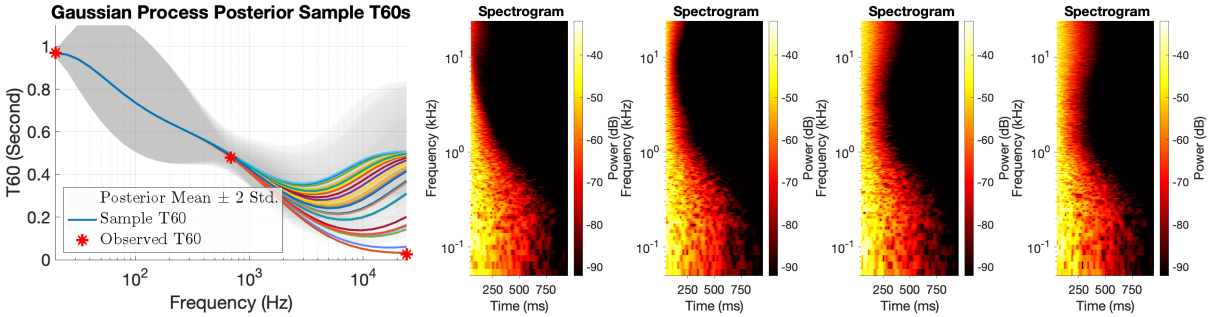


Fig. 5: The frequency responses of filters $\boldsymbol{g}$, fitted with fewer N taps to T_{60} targets via Eq. (17), exhibit amplified errors under exponentiation.

Directional Reverberation Generation: Consider a normally distributed real-valued random pressure field $\boldsymbol{P}$ represented by its first-order $(N_C = 1)$ SH expansion $\boldsymbol{C}$, and a GP distribution of direction dependent filters $\boldsymbol{g}(\theta,\phi)$. We specify a GP prior with the constant prior mean function $\overline{T}_{60}(\omega,\theta,\phi) = 0.5$ second in-place of Eq. (19), and non-stationary covariance of Eq. (21) with hyper-parameters $\alpha = 0.5$, $\beta = 0$, $\sigma = 0.25$, $\gamma = 0.6$, $\ell = 343$. Suppose we observe $\boldsymbol{y} = T_{60}(\underline{\boldsymbol{\omega}},\underline{\boldsymbol{\theta}},\underline{\boldsymbol{\phi}})$ at coordinates $\left\{\underline{\boldsymbol{\omega}},\left(\underline{\boldsymbol{\theta}},\underline{\boldsymbol{\phi}}\right)\right\} = \{\{0,693,24000\}\,\text{Hz} \times (90°,0°)\}$ following Eq. (18) as shown in Fig. 6a. Conditioning the GP on the observations, we jointly sample $T_{60}(\boldsymbol{\omega},\boldsymbol{\theta},\boldsymbol{\phi})$ at the predictive mean function $\boldsymbol{\mu}_*(\boldsymbol{X}_* = \{\boldsymbol{\omega} \times (\boldsymbol{\theta},\boldsymbol{\phi})\})$ in Eq. (20) over uniform frequencies $\boldsymbol{\omega} \in \mathbb{R}^N$ for $N = 16$, and uniform Fibonacci lattice [22] spherical coordinates $\boldsymbol{\theta},\boldsymbol{\phi} \in \mathbb{R}^{N_E}$, where $N_E = 25$ is the number of bases of a 4^{th}-order SH expansion. The sampled $T_{60}(\boldsymbol{\omega},\boldsymbol{\theta},\boldsymbol{\phi})$ encode our directional reverberation-time function as it veers from the observed $T_{60}(\underline{\omega},\underline{\theta},\underline{\phi})$ towards the prior mean $\overline{T}_{60}(\omega,\theta,\phi)$ for larger ω and $d(\theta,\phi,\underline{\theta},\underline{\phi})$. Thus, we realize the mean reverberation field of the GP by optimizing direction dependent exponentiating filters $\boldsymbol{g}(\boldsymbol{\theta},\boldsymbol{\phi}) \in \mathbb{R}^{N \times N_E}$ for design targets $T_{60}(\boldsymbol{\omega},\boldsymbol{\theta},\boldsymbol{\phi})$ on the spherical coordinate grid via Eq. (17) for $\tau = 1$.

Applying our directional filters $\boldsymbol{g}(\boldsymbol{\theta},\boldsymbol{\phi})$ to pressure field $\boldsymbol{P}$ in Eq. (13), we compute the exponentiating convolution pressure field matrix $\tilde{\boldsymbol{P}}$ via Algorithm 2 over the spherical grid $(\boldsymbol{\theta},\boldsymbol{\phi})$, and re-expand along SH bases to yield the directional reverberation expansion matrix $\tilde{\boldsymbol{E}} = \boldsymbol{Y}^{-1}(\boldsymbol{\theta},\boldsymbol{\phi})\tilde{\boldsymbol{P}}$. For cross-validation, we evaluate the expansion $\tilde{\boldsymbol{E}}$ at spherical coordinates on the horizontal plane outside the spherical grid $(\boldsymbol{\theta},\boldsymbol{\phi})$ as shown in Figs. 6b−6e. The resulting pressure/time spectrograms acquire the desired T_{60} profiles of our predictive mean function at azimuth angles that span the nearest to furthest spherical coordinates from the observation. Reducing the grid size to $N_E = 16$ introduced spatial aliasing in high-frequency for small T_{60} near $\phi = 0$.



(a) $T_{60}(\omega,\theta,\phi)$ at uniform spherical coordinates **(b)** $\phi = 0°$ **(c)** $\phi = 60°$ **(d)** $\phi = 120°$ **(e)** $\phi = 180°$

Fig. 6: Sampled T_{60} in Fig. 6a revert to the prior mean function $\overline{T}_{60}(\omega,\theta,\phi) = 0.5$ when further from the observed $T_{60}(\omega, 90°, 0°)$. Exponentiating convolutions are shown for varying ϕ on the horizontal plane $\theta = 90°$.

5 Conclusions

We presented several new methods for both augmenting and generating SH-RIRs with time-varying and direction dependent reverberation. First, we generalized the convolution operator for time-varying exponentiating filters and derived two fast algorithms for their efficient computation. We then extended our methods to SH transforms of multi-mic pressure fields by showing that exponentiating convolution commutes with all linear transforms. Next, we introduced direction dependent filtering and reverberation over finite-order SH bases using smooth filters in frequency and spherical coordinates. Our second-order cone program realized near minimum-phase filters from smooth $T_{60}(\omega,\theta,\phi)$ spectral-decay functions sampled from non-stationary GPs. Experiments validated our process for generating directional reverberation fields from Gaussian noise.

6 Appendix

Let $\boldsymbol{A} \in \mathbb{C}^{N_A \times N_H}$ be a linear transform matrix mapping N_H number of input to N_A output channels. Multiplying $\boldsymbol{A}$ by the results of the exponentiating convolution in Eq. (2) yields the matrix $\tilde{\boldsymbol{F}} = \boldsymbol{A}\boldsymbol{F} \in \mathbb{C}^{N_A \times MN}$ where $\boldsymbol{F} = \left[\boldsymbol{f}_1, \ldots, \boldsymbol{f}_{N_A}\right]^T \in \mathbb{R}^{N_H \times MN}$ is the row-vector matrix of exponentiating convolutions $\boldsymbol{f}_k = f(\boldsymbol{h}_k, \boldsymbol{g}) \in \mathbb{R}^{MN \times 1}$. The entrants of $\tilde{\boldsymbol{F}}$ are therefore given by

$$\begin{aligned} \tilde{F}_{in} &= \sum_{k=1}^{N_H} A_{ik} \boldsymbol{f}_k[n] \qquad \text{Post-convolution} \\ &= \sum_{k=1}^{N_H} \sum_{m=1}^{MN} A_{ik} h_k[n-m+1]\, g(n-m+1)[m]. \end{aligned} \tag{22}$$

Conversely, multiplying $\boldsymbol{A}$ by the multichannel matrix $\boldsymbol{H}$ in Eq. (7) before the exponentiating convolution yields $\tilde{\boldsymbol{H}} = \boldsymbol{A}\boldsymbol{H} \in \mathbb{C}^{N_A \times M}$ where $\tilde{\boldsymbol{h}}_i$ is the i^{th} row of $\tilde{\boldsymbol{H}}$, and $\tilde{H}_{in} = \sum_{k=1}^{N_H} A_{ik} h_k[n]$. Evaluating the exponentiating convolution over $\tilde{\boldsymbol{h}}_i$ in $f(\tilde{\boldsymbol{h}}_i, \boldsymbol{g})$ gives

$$\begin{aligned} f_i[n] &= \sum_{m=1}^{MN} \tilde{h}_i[n-m+1]\, g(n-m+1)[m] \quad \text{Pre-conv.} \\ &= \sum_{m=1}^{MN} \sum_{k=1}^{N_H} A_{ik} h_k[n-m+1]\, g(n-m+1)[m], \end{aligned} \tag{23}$$

which is equivalent to Eq. (22). Analogous substitutions of linear transformation orderings into recursive convolutions of Eq. (6) give the same equivalences.

In the case of linear transforms of the SH product operator $A(\boldsymbol{D})$ expressed by the transfer matrix $\boldsymbol{A}$, its n^{th} column has entrants specified along the SH bases with degree L bounded between $0 \le L \le L_D + L_C$, and order M bounded between $-L \le M \le L$ as follows:

$$\begin{aligned} A_L^M(n) &= \sum_{l=0}^{L_D} \sum_{k=\underline{k}}^{\overline{k}} \sqrt{\frac{(2l+1)(2k+1)}{4\pi(2L+1)}}\, c \begin{pmatrix} l & k & L \\ 0 & 0 & 0 \end{pmatrix} \\ &\quad \times \sum_{m=\underline{m}}^{\overline{m}} c \begin{pmatrix} l & k & L \\ m & M-m & M \end{pmatrix} D_l^m \mathbb{1}(n), \\ \mathbb{1}(n) &= \begin{cases} 1, & n = l^2 + M - m + k + 1 \\ 0, & \text{Otherwise} \end{cases}, \end{aligned} \tag{24}$$

where $c(*)$ are the Wigner's 3-j symbols [23], and $\mathbb{1}(n)$ is an indicator function of column index n. The summations indexes k are bounded between $\underline{k} = |L-l|$, $\overline{k} = \min(|L+l|, L_C)$, and m between $\underline{m} = \max(M-k, -l)$, $\overline{m} = \min(M+k, l)$.

References

[1] Müller, C., *Spherical harmonics*, Springer, 2006.

[2] Rafaely, B., *Fundamentals of spherical array processing*, volume 8, Springer, 2015.

[3] Jarrett, D. P., Habets, E. A., and Naylor, P. A., *Theory and applications of spherical microphone array processing*, volume 9, Springer, 2017.

[4] Zotter, F. and Frank, M., *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*, Springer, 2019.

[5] Wang, J., Zhang, J. A., Samarasinghe, P., and Abhayapala, T., "Time-domain wideband image source method for spherical microphone arrays," in *2023 IEEE 25th International Workshop on Multimedia Signal Processing (MMSP)*, pp. 1–6, IEEE, 2023.

[6] Xu, Z., Herzog, A., Lodermeyer, A., Habets, E. A., and Prinn, A. G., "Simulating room transfer functions between transducers mounted on audio devices using a modified image source method," *The Journal of the Acoustical Society of America*, 155(1), pp. 343–357, 2024.

[7] Luo, Y. and Kim, W., "Fast Source-Room-Receiver Acoustics Modeling," in *2020 28th European Signal Processing Conference (EUSIPCO)*, pp. 51–55, 2021.

[8] Hamilton, B., "Air absorption filtering method based on approximate green's function for Stokes' equation," in *2021 24th International Conference on Digital Audio Effects (DAFx)*, pp. 160–167, IEEE, 2021.

[9] Fagerström, J., Meyer-Kahlen, N., Schlecht, S. J., and Välimäki, V., "Dark velvet noise," in *International Conference on Digital Audio Effects*, pp. 192–199, DAFx, 2022.

[10] Kates, J. M. and Brandewie, E. J., "Adding air absorption to simulated room acoustic models," *The Journal of the Acoustical Society of America*, 148(5), 2020.

[11] Välimäki, V., Holm-Rasmussen, B., Alary, B., and Lehtonen, H.-M., "Late reverberation synthesis using filtered velvet noise," *Applied Sciences*, 7(5), p. 483, 2017.

[12] Zotter, F. and Frank, M., "All-round ambisonic panning and decoding," *Journal of the audio engineering society*, 60(10), pp. 807–820, 2012.

[13] Hold, C., Schlecht, S. J., Politis, A., and Pulkki, V., "Spatial filter bank in the spherical harmonic domain: Reconstruction and application," in *2021 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 2021.

[14] Rasmussen, C. E., "Gaussian processes in machine learning," in *Summer school on machine learning*, pp. 63–71, Springer, 2003.

[15] Pfaff, F., Kurz, G., and Hanebeck, U. D., "Filtering on the unit sphere using spherical harmonics," in *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pp. 124–130, IEEE, 2017.

[16] Luo, Y., "Spherical harmonic covariance and magnitude function encodings for beamformer design," *EURASIP Journal on Audio, Speech, and Music Processing*, 2021(1), p. 41, 2021.

[17] Jarosz, W., *Efficient Monte Carlo Methods for Light Transport in Scattering Media*, Ph.D. thesis, UC San Diego, 2008.

[18] Suda, R. and Takami, M., "A fast spherical harmonics transform algorithm," *Mathematics of computation*, 71(238), pp. 703–715, 2002.

[19] Paciorek, C. and Schervish, M., "Nonstationary covariance functions for Gaussian process regression," *Advances in neural information processing systems*, 16, 2003.

[20] Oppenheim, A. V., Schafer, R. W., and Buck, J. R., *Discrete-Time Signal Processing*, Prentice Hall, Upper Saddle River, NJ, 2nd edition, 1999.

[21] Pei, S.-C. and Lin, H.-S., "Minimum-phase FIR filter design using real cepstrum," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 53(10), pp. 1113–1117, 2006.

[22] Hannay, J. H. and Nye, J. F., "Fibonacci numerical integration on a sphere," *Journal of Physics A: Mathematical and General*, 37(48), pp. 11591–11601, 2004.

[23] Messiah, A., "Clebsch-Gordan (C.-G.) coefficients and'3j'symbols," *Appendix CI in Quantum Mechanics*, 2, pp. 1054–1060, 1962.