arXiv:2602.01758v1 [eess.AS] 2 Feb 2026

时域耳蜗传输线模型的短波导纳校正

François Deloche *[1]、Morgan Thienpont[1], 和 Sarah Verhulst[1]

[1]Hearing Technology @ WAVES，根特大学信息技术系，比利时根特

2026年2月

摘要

在时域中实现的传输线（TL）模型能够高效模拟基底膜（BM）对瞬态或非平稳声音的位移响应。按照设计，传输线模型适用于对行波进行一维（1-D）描述，但耳蜗的真实结构也会产生高维效应。这些效应包括基底膜附近的压力聚焦和横向黏性阻尼，在短波区域二者都会增强。这两种效应取决于波长，在频域中更容易表达。本文采用自回归滤波和回归技术，提出一种数值校正，以便在时域中对基底膜导纳进行校正，从而补偿二维效应。该校正是实现适用于沙鼠耳蜗生理特性的传输线模型所必需的。该模型以可变阻尼的形式纳入了瞬时非线性，但最初在声级升高时表现出的压缩不足。我们将这一局限归因于一维非线性传输线模型中假设的增益与频率选择性之间的强耦合；而小型哺乳动物的耳蜗频率选择性仅适度依赖于声级。我们在沙鼠模型中实现了校正因子，并通过反馈回路使其具有声级依赖性。更新后的模型在一定程度上实现了频率选择性与增益的解耦，额外提供了5分贝增益，并将压缩区间的声级范围扩展了10分贝。我们通过两个关键特点讨论这项工作的意义：结合解析方法和回归方法来表征基底膜导纳，以及将瞬时非线性与非瞬时非线性相结合。

I 引言

耳蜗宏观力学可简化为单个自由度，例如基底膜（BM）的位移，并可通过传输线（TL）模型加以描述，该模型能够捕捉耳蜗行波的基本特征。将TL模型纳入能够重现听觉通路各个阶段（即内毛细胞、听觉突触与听神经、脑干核团）的处理流程后（Verhulst 等人（2018）），它们在生成与听觉外周相关的诱发响应方面，能够与现象学模型相媲美（Saremi等，2016；Vecchi等，2022）。TL模型的优势包括仿真耳声发射以及重现瞬时非线性。时域仿真需要精细的时空采样，但可通过高效的数值方案加快计算（Altoè等，2014）；更有效的方法是训练深度神经网络，以重现声音输入与基底膜输出之间的映射（Baby等，2021）。

TL模型能够有效捕捉声音作用下沿基底膜传播的行波，但其一维（1-D）特性可能使整合跨越多个维度的效应变得复杂。其中一些效应可以隐含地整合到对行波的一维描述中。例如，“尺度对称性”这一性质将基底膜导纳表示为归一化

*电子邮件：francois.deloche@polytechnique.org

频率的函数，与耳蜗的锥形几何结构有着形式上的联系（Shera 和 Zweig 1991）。这一性质确保驱动压在达到峰值之前以恒定幅值传播，类似于平面波（Altoè 和 Shera 2020）。另一个需要考虑的方面是 TL 模型的稳定性，因为主动耳蜗中每个特征位置的基底侧区域都存在负阻尼，使耳蜗模型容易出现不稳定。在二维或三维模型中，可以通过添加一个取决于流体速度横向梯度的粘滞项来维持稳定性（Sisto 等人，2021年、2023年）。随着波接近峰值区域，与基底膜速度相比，该粘滞项的幅值会增大：波的阻尼迅速增强，从而自然地抑制负阻尼区域。时域 TL 模型的情况则更为复杂。Zweig 提出了一种巧妙的变通方法（Zweig 1991），Shera 随后对其加以发展（Shera，2001）：在基底膜上添加作用缓慢、具有时延的刚度，将负阻尼区域限制在特征位置的基底侧区域。该方案被整合进一个全面的听觉外周计算框架中（Verhulst 等人，2012年、2015年、2018年）。相应的 TL 模型称为 V-1D 模型，它是本文所提出模型的基础。

尽管 V-1D模型中的基底膜导纳包含一个慢反馈项，但它仍然基于经过修改的谐振子。因此，增益与频率选择性紧密耦合，类似于经典的二阶滤波器：调谐范围较窄的基底膜导纳会产生峰值较高的频率响应，而较宽的频率响应则对应较低的增益。对于模拟人类听觉响应而言，这一特性并不是关键问题，因为人类的耳蜗滤波器被认为明显比小型哺乳动物的滤波器更尖锐（Shera 等人；Verschooten 等人 2018；Shera 和 Charaziak 2019）。这种更尖锐的调谐与耳蜗响应的宽广压缩范围非常吻合。对于再现常见实验动物听觉响应的模型而言，这一问题则更加明显；这些动物尤其受到关注，因为可以将模型与更广泛的生理数据进行对比，而获取这些数据需要采用侵入性技术。对于其中许多动物（例如沙鼠）而言，在耳蜗处于最敏感状态的低声级下，基底膜速度频率响应的峰值较高，但相对宽广（Fallah 等人 2021；He 等人（2022））。考察多个声级时，压缩范围很大，但调谐变宽却十分有限。这一表面上的不一致给近期将 V-1D模型调整用于小型啮齿动物，特别是沙鼠和小鼠，带来了挑战（Thienpont 等人 2024）。在我们最初尝试匹配沙鼠耳蜗响应时，为实现所需的响应增益变化，模型缺少约 10 分贝的压缩。

在耳蜗的二维或三维模型中，波峰幅值的部分成因可归结为短波区域发生的压力聚焦现象。这种效应源于波长缩短时压力场特征高度的降低（Reichenbach 和 Hudspeth，2014年）。随着接近波峰，特征高度逐渐减小，压力便会在基底膜周围的薄层中“聚焦”。这会使驱动基底膜速度的压力获得一个正增益因子。这种现象是被动的，但如果耳蜗压缩非线性伴随着波长增加，它在最活跃状态下的影响会更强（Sisto 等人，2021）。在这种情况下，压力聚焦会扩大耳蜗响应增益的范围，同时对频率选择性的影响有限。

本文采用若干数值方法，在针对沙鼠生理特征定制的传输线模型中模拟压力聚焦带来的增益补偿。该校正基于 Sisto 等人的研究，该研究将二维流体效应（黏性阻尼和压力聚焦）纳入频域传输线模型（Sisto 等人，2021）。我们通过自回归滤波和回归方法在时域实现了这一校正。除了本文介绍的特定任务改进外，这种新方法还可能适用于多种需要在传输线模型中考虑波长相关效应的情形，而这一直是时域实现面临的挑战。

II 方法

在方法部分，我们采用以下命名规则来指代本研究旨在结合的两种传输线（TL）模型：

- **V-1D模型**指的是 Verhulst 等人（Verhulst

等人（2018）所描述的时域传输线实现，该实现基于 Zweig 和 Shera 对基底膜导纳的表征（Zweig 1991；Shera，2001）。

- **S-2D 模型**指的是 Sisto 等提出的压力波频域 WKB 近似（Sisto 等人，2021）。该模型纳入了两种二维流体动力学效应：压力聚焦和黏性阻尼，并作为本文所引入的短波导纳校正的参考。

我们提出了第三种传输线模型，即 $V^\star$ **模型**。该模型主要基于 V-1D模型的框架，同时融入了 S-2D 模型的部分要素。

方法部分分为两部分。前三个小节回顾 V-1D模型和 S-2D 模型的关键要素，为 $V^\star$ 模型建立共同框架。最后一个小节介绍短波导纳校正的实现。

II.1 压力聚焦引起的短波增益增强（S-2D 模型）

我们首先介绍 S-2D 模型，因为该模型阐明了传输线方程与耳蜗行波二维物理模型之间的联系。这些方程综合考虑了压力聚焦和作用于基底膜的流体黏性应力。更多详情请参见原论文（Sisto 等人，2021）。

在 S-2D 模型中，上、下耳蜗管道被视为两个高度为H 、由基底膜隔开的二维矩形盒体。坐标 x 和 z 分别对应纵向和横向。我们将 $p_d = p(z) - p(-z)$ 定义为基底膜上方的耳蜗管道与鼓阶之间的压力差。 $U = \int_0^H u_x$ dz 是上方耳蜗管道中的体积速度，而 p_d 是沿横向平均的差压。 $\overline{}$

描述压力波的“类似传输线”方程为：

$$\frac{\partial \overline{p_d}}{\partial x} = -Z_f\, U \tag{1a}$$

$$\frac{\partial U}{\partial x} = v_{BM} = Y_{BM}\, p_d(z = 0^+), \tag{1b}$$

其中$Z_f = \frac{2j\omega\rho}{H}$，而 Y_{BM} 是基底膜导纳，用于描述基底膜速度与施加在基底膜两侧的压力差之间的关系。

这些方程几乎化简为一维情形，但在方程 1a 中，压力 p_d 表示沿横向的平均值，而在方程 1b 中， p_d 是在 $z = 0^+$ 处求得的，该位置对应基底膜上表面。引入 $\alpha = p_d(z = 0^+)/p_d$（即“压力聚焦”因子）后，第二个方程可改写为：

$$\frac{\partial U}{\partial x} = Y_{BM}\, \alpha\, \overline{p_d}\ . \tag{2}$$

假设 $Y_{BM}\,\alpha$ 是位置 x 和角频率 ω 的已知函数，则可使用一维形式体系。结合方程 1a 和 2，可得波动方程：

$$\frac{\partial^2 \overline{p_d}}{\partial^2 x} = -\kappa^2 \overline{p_d}\ , \tag{3}$$

其中 $\kappa^2 = Z_f Y_{BM} \alpha$。

将 WKB近似应用于二维箱模型，可得压力聚焦因子的以下公式：

$$\alpha = \frac{\kappa H}{\tanh \kappa H} \ , \tag{4}$$

从而得到色散关系（Reichenbach 和 Hudspeth，2014年）：

$$\kappa \tanh(\kappa H) = 2j\rho\omega \, Y_{BM} \, . \tag{5}$$

在 Sisto 等的研究中，这种方法被扩展为纳入作用于基底膜（BM）的流体黏性应力的影响。黏性应力与 $z = 0$处（横向）流体速度的横向梯度成正比。在经典势流描述中，后者与沿 z的压力二阶导数有关，而 v_{BM}则与一阶导数有关。WKB近似给出了二阶导数与一阶导数之比的 α/H 。因此，作用于基底膜的黏性应力（Sisto 等对此作了详细推导）为:

$$S = -4b\mu \, \frac{\alpha}{H} v_{BM}, \tag{6}$$

其中， μ 是流体的动力黏度， $b = 2.5$ 是用于考虑流体流动中非理想部分的经验因子（Sisto 等人，2021）[1]。基底膜导纳根据简谐振子模型推导得出，并经过修正以纳入黏性应力:

$$Y_{BM} = -\frac{j\omega}{\sigma_{BM}} \left[-\omega^2 + j\omega \, (\Gamma + \frac{4\alpha b\mu}{\sigma_{BM} H}) + \omega_{bm}^2 \right]^{-1} \ . \tag{7}$$

在此公式中， ω_{BM} 是基底膜的特征频率， σ_{BM} 是基底膜的有效质量密度。 Γ 是标准阻尼项，可进一步分解为主动分量和被动分量：$\Gamma = \Gamma_a + \Gamma_p$，其中 $\Gamma_p = \omega_{BM}$且 $\Gamma_a = -G\omega_{BM}$。变量 G表示主动过程的强度，它会产生负阻尼。

更新后的基底膜导纳得出以下色散关系:

$$\kappa \tanh(\kappa H) = \frac{2\omega^2 \rho}{\sigma_{BM}} \left[\omega_{bm}^2 - \omega^2 + j\omega \, (\Gamma + \frac{4\alpha b\mu}{\sigma_{BM} H}) \right]^{-1} \ . \tag{8}$$

这一关系使 Y_{BM} 独立于 κ的情况变得更加复杂，因为取决于 κ)的因子 α出现在方程的右侧。为计算 κ，我们采用 Sisto 等描述的递归过程。收敛后，使用式4计算压力聚焦因子 α 。为撰写本文，我们还引入 β 因子，用以描述α 相对于参考值的变化:

$$\beta = \alpha/\alpha_0 \, , \tag{9}$$

其中， α_0 是在参考水平下计算的压力聚焦因子。我们计算中使用的 S-2D 参数与 Sisto 等采用的参数基本相同（见表1），但我们调整了H 的值，以反映沙鼠耳蜗在基底与耳蜗顶端中点处的高度（$CF = 4$千赫）。

II.2 时域实现：基线（V-1D模型）

在介绍整合压力聚焦因子的策略之前，我们先回顾一下作为时域实现基础的 V-1D 传输线模型的一些要素。V-1D 模型由$N = 1000$ 级联的若干段组成，用于表示从基底到耳蜗顶端的耳蜗分隔 N 各段。

[1]请注意，较新的一项研究（Sisto 等，2023年）未纳入经验因子 $b = 2.5$ ；此处我们保留该因子，以与2021年论文保持一致。

每段都包括一个串联阻抗 Z_{s_n} 和一个并联导纳 Y_{p_n} （Zweig 1991；Verhulst 等人 2012、2018）：

$$Z_{s_n}(j\omega) = M_{s_n} j\omega \ , \tag{10}$$

$$Y_{p_n}(j\omega) = \frac{j\omega}{M_{p_n}} \left[\omega_n^2 - \omega^2 + \delta_n j\omega\omega_n - \varrho_n e^{-j\psi_n \omega/\omega_n}\omega_n^2\right]^{-1} \ . \tag{11}$$

n 从 0（基底）到 $N-1$ （耳蜗顶端）为传输线段编编号。质量参数按渐缩导波管的规律缩放，其中刚度与特征频率成正比（Zweig 1991；Shera and Zweig 1991；Altoè and Shera 2020）： $M_{s_n} = \frac{\omega_0}{\omega_n} M_{s_0}$，以及 $M_{p_n} = \frac{\omega_0}{\omega_n} M_{p_0}$。该模型被视为一组耦合微分方程，并采用龙格－库塔法（RK）4(5)这一有限差分法进行数值求解。 $\varrho_n e^{-j\psi_n \omega/\omega_n}\omega_n^2$ 对应一个慢反馈项，其刚度为 ϱ_n ，延迟为 ψ_n，并通过样条插值将其纳入 RK4(5)格式（Altoè等，2014）。

V-1D模型包含取决于基底膜速度幅值的瞬时非线性。具体而言，三元组 $(\varrho_n, \delta_n, \psi_n)$ 会随时间更新，以改变基底膜阻尼，同时确保相应的（双）极点轨迹与基底膜速度零交叉点近乎不变这一特性保持一致（Shera，2001）。关于如何根据基底膜速度确定这三个参数的公式，可参见此前的研究（Verhulst 等人（2018））。s域中双极点到虚轴的距离决定了局部基底膜脉冲响应的阻尼。在低声级下，该距离最小；在响应呈压缩特性的声级范围内，它会逐渐增大并趋于饱和值（Verhulst 等人（2018））。这种行为模拟了耳蜗调谐和响应增益对声级的依赖性。最终，模型中的非线性仅由初始状态以及参数 a 决定；该参数控制响应随声级升高而产生的压缩程度（分贝/分贝）。初始状态反映了低声级下的耳蜗，此时基底膜响应的调谐最尖锐，灵敏度也最高（Ruggero et al. 1997）。对于动物模型，压缩参数可由基底膜速度增长函数推导得出，即对数刻度下所表示的增长函数的斜率（Thienpont 等人（2024））。

本研究使用的模型版本经过校准，以符合沙鼠生理特征（参数见表1）。该模型与 Thienpont 等（2024）所述的沙鼠模型相同，但调整了频率选择性参数，使其反映基底膜调谐，而非听神经调谐。我们这样做是为了便于将模型与基底膜振动数据进行比较。V-1D模型中的极点根据品质因数 Q_{ERB} 与中心频率之间的幂律关系设定（Verhulst 等人（2018））。在 Thienpont 等（2024）中，该幂律关系与听神经数据（Müller 1996）相匹配。本研究使用了相同的幂律关系，但将 Q_{ERB} 除以 1.9，以反映基底膜调谐。这一因子对应于一项近期沙鼠耳蜗频率选择性研究报告的听神经调谐与基底膜调谐之间的品质因数比（Charaziak and Shera 2021）。

II.3 两种 TL 模型之间的联系

到目前为止，我们介绍了两种现有的 TL 模型（V-1D 和 S-2D）。在方法部分的这一阶段，有必要指出它们的相似之处和差异，并强调在本文其余部分中，每种模型都有各自不同的用途。这两种模型都旨在描述耳蜗行波及其随声级变化的非线性行为。它们均基于压力和速度的耦合方程：对于 S-2D 模型，方程以偏微分方程的形式给出（方程 1a 至 2）；对于 V-1D 模型，则以阻抗系统的形式给出（方程 10 和 11）。因此可以看出，这两种模型在形式上有许多相似之处，但为避免混淆各自的参数，采用了不同的符号。

然而，这两个模型在几个方面存在差异。S-2D 模型基于物理原理。该模型在频域中更易于进行数值求解，但即使在频域中，基底膜阻抗也没有闭式表达式。相比之下，V-1D模型专门设计用于提供

相对简单的基底膜导纳表达式，并再现过零点不变性等时域特性（Shera，2001）。因此，将 V-1D模型作为时域实现的基础更为自然。V-1D模型导纳中的慢延迟反馈项与耳蜗中任何已知的生理机制都没有关联，但它提供了必要的频率调制，使频域中的负阻区域和正阻区域交替出现，从而确保数值稳定性。

这两个模型在模拟行波响应时也表现出差异，尤其是对波峰区域的表示方式。在 V-1D模型中，波峰出现在刚度主导区域与质量主导区域的交界处；而在 S-2D 模型中，波峰出现在更靠基底侧的位置（即刚度主导区域内）。讨论部分将进一步强调这一差异。这里我们只指出，这会导致两个模型对局部基底膜共振频率产生不同的解释。在 V-1D模型中，纯音响应的行波波峰与特征角频率 $\omega_n = 2\pi f$所在区域重合。换言之，在低声级下，特征频率 CF 与最佳频率 BF 相同（Shera，2001）。在 S-2D 模型中，行波速度减慢，并在到达由质量与刚度之比所定义的特征位置之前达到波峰（Sisto 等人，2021年；Sisto 和 Moleti，2024年），因此波峰位置处的频率为 $BF < CF$。在实际应用中，通过对 ω_{BM} 应用经验仿射关系来消除这种差异，使两个模型的最佳频率一致。基于 S-2D 模型的 WKB 模拟，我们确定 $\omega_{BM} = 1.2\,\omega_n + 1500\,\text{Hz}$，其中 ω_{BM}是 S-2D 模型的特征角频率，而最佳频率等于 $\omega_n/(2\pi)$。在计算压力聚焦增益时，利用这一关系为每个传输线截面赋予 ω_{BM} 的值。与 V-1D模型一样，位置-频率映射由 Greenwood函数设定：

$$\omega_n = 2\pi A_1(10^{-A_2 x} - B)\ , \tag{12}$$

其中，x 是距基底的距离（单位：米）。S-2D 模型仅用于计算压力聚焦因子，该因子是局部属性，因此，为了实现以下小节中所述的方法，我们无需为 S-2D 设置音调定位映射。

II.4 导纳校正的实现（ $V^\star$ 模型）

图 1 a nd 2 列出实现的主要步骤，分为设置阶段和运行时。

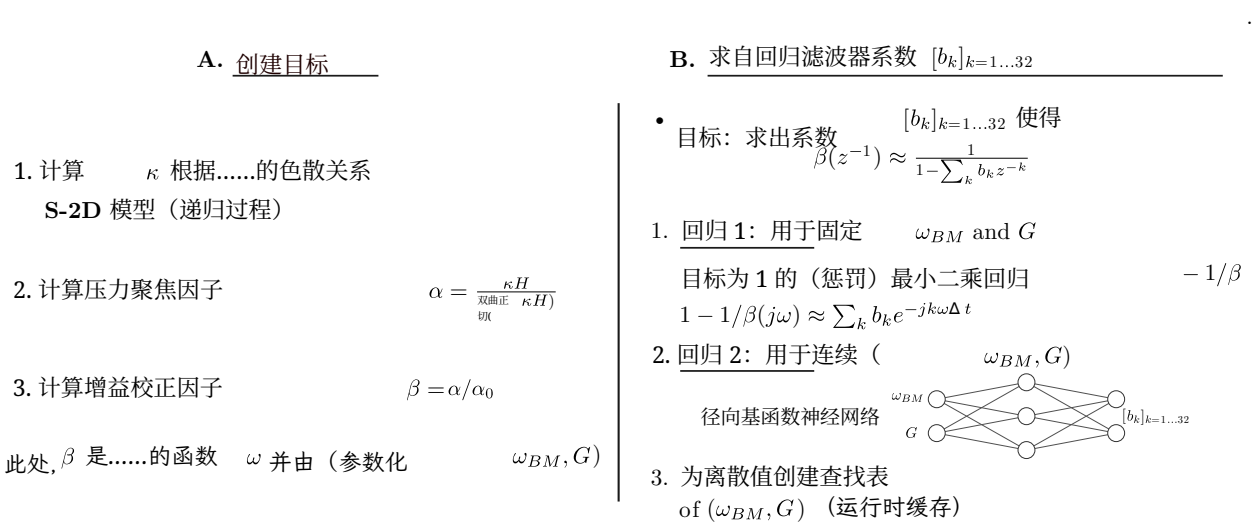


图 1：导纳校正设置的实现策略。

总体策略如下：

运行时

A. 基底膜导纳中的积分

改编自Diependaal等人 .

/ Altoè et al. (**V-1D** model)

1. 计算:

$$Q_n(t) = \sum_{k=1}^{32} b_{n,k}\ (\dot{v}_n[t - k\Delta t] + g_n[t - k\Delta t])$$

2. 求解**A q = g - Q**

3. 计算 $\dot{v}_n[t] =$ $q_n[t] - g_n[t] + Q_n[t]$.

B. 系数更新 [$b_{n,k}]_{k=1\ldots32}$:

1. 计算 （输出强度），使用一阶滤波器

$v_n^2 \rightarrow H(z^{-1}) \rightarrow < I_n >$ $\quad H(z^{-1}) = \frac{c_{n,0} + c_{n,1} z^{-1}}{1 + d_{n,1} z^{-1}}$

2. 查找 G_n 使用线性 在.....之间插值 G_{min} and G_{max} 取决于 $< I_n >$ 以分贝为单位

3. 更新 [$b_{n,k}]_{k=1\ldots}$ 该截面的 32 n 使用:

$(\omega_{BM}, G) = (1\ 2.\ \omega_n + 1500, G_n)$

图2：运行时的实现策略（运行仿真时）。

1. 创建函数 $\beta = \beta(\omega_{BM}, G)$，根据 S-2D 模型计算并返回应用于基底膜导纳的校正因子。该函数的输入是基底膜特征频率和“主动过程”的强度。输出是在频域中定义的向量。2. 确定自回归系数，以将 β 近似为全极点滤波器的频率响应。滤波器系数记为 $[b_k]_{k=1\cdots32}$，且也取决于 ω_{BM} 和 G。此步骤包含两种依次执行的回归方法。3. 更新时域方程，将由系数 $[b_{n,k}]_{k=1\cdots32}$ 定义的递归滤波器 β 应用于 V-1D 基底膜导纳。4. 仿真过程中：随时间更新参数 $[b_{n,k}]_{k=1\cdots32}$，以反映主动过程的变化；其强度随响应强度而变化。

步骤 1 和 2（设置阶段）只需运行一次，即可找到定义递归滤波器的回归参数。运行仿真时执行步骤 3 和 4。以下各小节将详细介绍每个步骤。

II.4.1 自回归滤波器系数的估计

我们的目标是通过引入根据 S-2D 模型计算得到的压力聚焦因子 α（式 4），补上 V-1D 模型中缺失的增益。不过，由于 α 可能取较大值，难以直接纳入时域实现，因此我们改用先前引入的归一化 β 因子（式 9：$\beta = \alpha/\alpha_0$）。 α_0 是以活跃过程强度的参考值 $G = G_{REF}$ 计算得到的压力聚焦因子。 G_{REF} 取 G （$[G_{\min}, G_{\max}]$）取值范围的中点，以使 β 尽可能接近 1。为简洁起见，我们省略了 β的完整依赖关系，其完整形式应写为 $\beta(\omega_{BM}, G, \omega)$。我们称 β为增益增强或校正因子，尽管在实际应用中其幅值可能小于 1。还要注意， β 的输出为复数，即 β 不仅由幅值定义，也由相位定义。

对于给定的一对 (ω_{BM}, G)，我们将 β 近似为全极点滤波器的频率响应。具体而言，我们希望找到系数 $[b_k]_{k=1\cdots32}$，使得:

$$\beta(z^{-1}) \approx \frac{1}{1 - \sum_{k=1}^{32} b_k z^{-k}}. \tag{13}$$

我们采用 32 个系数，在精度和计算成本之间取得平衡；不过，该方法具有通用性，也可以使用其他数量的系数。在频域中，我们还可以写成

$$1 - 1/\beta(j\omega) \approx \sum_{k=1}^{32} b_k e^{-jk\omega\Delta t}\ . \tag{14}$$

表1：模型和回归参数。**上：** S-2D 模型的参数与（Sisto 等人，2021）（采用 $\mu = 10\mu_w$ 的 WKB 模拟）中的参数相同，粗体显示的参数除外。**中：** V-1D 和 $V^\star$ 模型的主要参数。有关模型参数化的更多详情，请参阅 Thienpont 等（2024）。

S-2D 模型（计算 α 和 β）	
$\mathbf{H}$	320 µm
σ_{BM}	0.06 千克/米^2
ρ	1.0×10^3千克/米^3
$\mu = 10\mu_w$	7.0 mPa s
$[\mathbf{G}_{\min}, \mathbf{G}_{\max}]$	[0, 1。3]
$\mathbf{G_{REF}}$	**0.7**
V-1D / V* 模型（时域实现）	
f_s （采样频率）	200 千赫
N （分段数）	1000
H	320 µm
L（耳蜗长度）	12.1 mm
Greenwood A_1	50 216 Hz
Greenwood A_2	181.034 米$^{-1}$
Greenwood B	140 赫兹
Q_{ERB} ($f_0 = 1$ 千赫)	$1.45 \cdot (f/f_0)^{0.58}$
a (压缩斜率)	0.45 分贝/分贝
回归	
λ_1 （第一个最小二乘误差回归惩罚项）	1
λ_2 （第二个最小二乘误差回归惩罚项）	0.3
径向基函数神经网络参数	在正文中。

我们使用此公式求出系数 b_k。我们在等间隔的频率上对等式进行离散化，将等式两边分别表示为频域中的向量。随后，我们确定使这两个向量之间的最小二乘误差（LSE）最小的系数 b_k 。我们只计算频率不超过 $\omega_{cut} = 1.3\,\omega_{BM}$ 时的 LSE，因为传输线会滤除更高的频率。不过，我们还加入了两个惩罚项，以确保：a) 系数 $[b_k]_{k=1\cdots32}$ 保持较小；b) 对于高于 ω_{cut} 的频率，函数$\hat{\beta}(j\omega) = \left[1 - \sum_{k=1}^{32} b_k e^{-jk\omega\Delta t}\right]^{-1}$ 的虚部相对于实部较小或为负。

第一个条件通过对 $[b_k]_{k=1\cdots32}$ 的平方和施加惩罚来实现，如同岭回归。第二个条件则通过对超过 ω_{cut} 的 $\max\left(0, \mathrm{Im}(\hat{\beta})/\mathrm{Re}(\hat{\beta})\right)$ 施加惩罚来控制，其中最大值算子逐元素应用。该惩罚可防止传输线模型的单色解（局部正比于 $\exp(-j\kappa x)$）因 κ 中存在正虚部而在高频下变得不稳定。

附录A进一步介绍了用于确定 $[b_k]$ 系数的最小二乘回归，包括代价函数的完整形式。我们还在附录B中介绍了一种方法，用于确保校正因子在低频时的增益为零。采取这一预防措施是为了确保低频波以平面波的形式传播至其峰值（即具有恒定的特性导纳）Altoè 和 Shera（2020），因为反之可能导致伪影行为（Shera 和 Zweig 1991）。

上述回归方法适用于给定的一对 (ω_{BM}, G)。为了在允许 (ω_{BM}, G) 变化时对 $[b_k]$ 系数进行连续插值，我们采用了第二种回归方法，即径向基函数神经网络（RBF-NN）。网络的输入是一个元组 (ω_{BM}, G)，输出是一组估计系数 $[b_k]_{k=1\cdots32}$，根据公式 13，这些系数近似于 β。以下

句将进一步介绍 RBF-NN 及其训练方法。首先对输入进行归一化，使其取值介于 0 和 1 之间。网络包含 360 个径向基中心（18×20），每个中心都对应一个标准差为 0.04 的高斯核。核在第一维度上采用“局部线性”形式，即乘以一个含有可变参数 ω_{BM} 的仿射函数。RBF-NN 使用随机梯度下降法，在随机生成的成对数据 (ω_{BM}, G) 上进行训练。定义代价函数时考虑了两种策略。第一种策略使用第一种方法（惩罚线性回归）生成目标 $[b_k]$，并最小化目标与 RBF-NN 输出之间的欧几里得距离。第二种策略直接最小化 LSE 代价函数（公式 14），跳过通过第一种方法生成 $[b_k]$ 系数的步骤。我们使用第一种策略对 RBF-NN 进行了 1,000 步训练（学习率：0.1，动量：0.8），随后又进行了 1,000 步训练，使用第二种策略微调参数（学习率：1e-3，动量：0.9）。

在模拟刺激响应中的行波时，使用径向基函数神经网络实时生成 β 系数可能会带来较高的计算开销。为减少计算时间，我们允许预先计算 beta 系数，并将其存储在大小为 1000×30的查找表中，其中 $N = 1000$ 为通道数，30 为覆盖 G 值范围的离散步数。该查找表只需创建一次，之后即可在所有仿真中重复使用。

II.4.2 在 V-1D模型中集成新滤波器

注：在本小节中，我们在系数 $[b_{n,k}]$ 上添加下标 n，以强调滤波器系数取决于由 n。索引的传输线截面。

传输线方程（方程 11）中的并联导纳现在包含增益校正因子 $\beta(j\omega)$，因此需要相应地更新时域方程。我们按照 Altoè 等（2014）的做法，引入以下变量：

$$g_n(t) = \omega_n^2 y_n(t) + \delta_n \omega_n v_n(t) - \varrho_n \omega_n^2 y_n(t - \tau_n),$$

$$q_n(t) = \frac{1}{L_{BM} M_{p_n}} p_n(t),$$

其中，$\tau_n = \psi_n / \omega_n$ 和 L_{BM} 是基底膜宽度。包含压力聚焦因子的运动定律为：

$$\dot{q}_n = (\dot{v}_n + g_n) \cdot (1 - \sum_{k=1}^{32} b_{n,k} z^{-k}) \,. \tag{15}$$

将当前和过去的样本分开，可得：

$$\begin{aligned} \dot{v}_n[t] &= q_n[t] - g_n[t] + Q_n[t], \text{where} \\ Q_n[t] &= \sum_{k=1}^{32} b_{n,k} \left(\dot{v}_n[t - k\Delta t] + g_n[t - k\Delta t] \right) \,. \end{aligned} \tag{16}$$

方程16与Diependaal等（1987）和Altoè 等（2014）中的更新方程非常相似，但通过新变量$Q_n(t)$引入了对过去样本的依赖。第二组方程由 Δp 和 $\dot{v_n}(t)$ 之间的关系推导得出（Diependaal等人，1987），可写为：

$$\mathbf{A}\mathbf{q} = \mathbf{g} - \mathbf{Q} \tag{17}$$

其中**A** 是一个三对角矩阵（例如参见 Verhulst 等人（2012）中的公式 2），而**q**、**g**和**Q** 分别是 $[q_n]$、$[g_n]$和 $[Q_n]$的向量形式。

基底膜运动的模拟采用与 Diependaal 等人（1987）相同的方法，即在每个时间步先求解公式 17，再求解公式 16。为提高准确性和效率，V-1D模型采用 RK4(5) 自适应数值方案（Altoè等，2014），该方案需要使用三次样条对过去的值进行插值，以处理长延迟反馈项。我们采用相同的策略，在每个内部步骤估算 Q_n。一个难点是，Q_n 包含一项取决于 $\dot{v}_n[t-\Delta t]+g_n[t-\Delta t]$（对于 $k=1$）的项，由于在时间步 t 时 $\dot{v}_n[t+\Delta t]$ 的值尚未知，因此无法使用三次样条进行插值。为解决这一问题，我们将Q_n 分为两部分：第一项 $b_{n,\ 1}\left(\dot{v}_n[t-\Delta t]+g_n[t-\Delta t]\right)$，使用 $t-2\Delta t$、$t-\Delta t$ 和 t 处的点，通过拉格朗日二次插值进行估算。第二项 $\sum_{k=2}^{32} b_{n,\ k}\left(\dot{v}_n[t-k\Delta t]+g_n[t-k\Delta t]\right)$，则使用相同的点和 $t+\Delta t$，通过 Catmull-Rom 三次样条进行插值。

II.4.3 仿真期间更新滤波器系数

校正增益因子 β 取决于主动过程的强度（S-2D 模型中的变量 G），而主动过程的强度会随时间根据响应强度而变化。对于每个耳蜗分段，我们使用一阶递归滤波器计算强度的短时平均值：

$$\langle I_n\rangle[t]=c_{n,0}\,v_n^2[t]+c_{n,1}\,v_n^2[t-\Delta t]-d_{n,1}\,\langle I_n\rangle[t-\Delta t]\ . \tag{18}$$

系数 $[c_n]$、$[d_n]$ 经过选择，用于实现截止频率为 $f_n/2=\omega_n/(4\pi)$的一阶低通巴特沃斯滤波器。该截止频率可确保包络 $\langle I_n\rangle(t)$ 平滑，同时仍能对信号强度的变化迅速作出响应。

$V^\star$ 模型保留了 V-1D 模型中依赖瞬时速度的非线性，并将其划分为三个不同区间：低于第一个转折点 v_{knee1}时为线性区间，介于 v_{knee1} 与第二个转折点 v_{knee2}之间为压缩区间，高于 v_{knee2}时则为第二个线性区间。压缩区间内基底膜导纳（双）极点值的变化轨迹由恒定的压缩斜率 a决定，并采用双曲插值（Verhulst 等人（2018））。

$V^\star$ 与 V-1D 模型在确定转变点的方式上略有不同。膝点通过对上述两个线性区间下的纯音响应进行仿真来确定。对于低声级和高声级区间，相应的极点分别称为起始极点和饱和极点。对于沙鼠模型，我们绘制了 v_{BM} 在特征频率为 =20 千赫的位置处的增长函数。由于这两个区间均为线性，当以纯音级（两者的单位均为分贝）为横轴绘制 v_{BM} 时，增长函数呈直线。利用起始极点得到的增长函数来确定 v_{knee1}，对应于 $I_{knee1\ \ dB}$，即标志着压缩起始点的声级。第二个转变点 v_{knee2} 则通过使用饱和极点进行仿真来确定：求增长函数与方程为 $v_{dB}=v_{knee1\ \ dB}+a(I_{dB}-I_{knee1\ \ dB})$ 的直线的交点，其中 a 为压缩斜率。为校准 $V^\star$ 模型，我们采用相同的步骤，但在起始极点的情况下，将压力聚焦因子设为其最大值（$G=G_{\max}$）；在饱和极点的情况下，则设为其最小值（$G=G_{min}$）。压缩非线性的校准将在结果部分末尾进行说明。

一旦设定了转变点，取决于 G 的压力聚焦因子便可动态更新。在压缩区间内，G 由 $G_{\min}$ 和 $G_{\max}$ 之间的线性插值确定，其中速度值以分贝表示：

$$G(t)=G_{\max}+\frac{\langle I_n\rangle_{dB}(t)-6\,dB-v_{knee1_dB}}{v_{knee2_dB}-v_{knee1_dB}}(G_{\min}-G_{\max})\ .$$

−6分贝因子是一种经验因子，用于使 $\langle I_n\rangle$ 与均方根值$v_{knee1,2}$保持一致。在沙鼠模型的仿真中，系数 $[b_k]$ 会根据此公式每隔 6 个时间步更新一次（即在 $f_s=200$千赫下每隔 0.03 毫秒更新一次）。

III 结果

III.1 V-1D模型中的压缩响应

我们首先介绍最初的 V-1D 模型在适配沙鼠生理特征时压缩不足的问题。图 3A 展示了纯音刺激下的基底膜速度增长。仿真结果正确呈现了一个压缩区间，其两侧分别是低声级和高声级下的线性区间。然而，发生压缩的声级范围不足 30 分贝。针对人类听觉定制的模型版本（Verhulst 等人（2018））的压缩区间覆盖 40 分贝，且压缩斜率更大。图 3 所示的频率响应有助于进一步理解这些差异的成因。这些频率响应是通过仿真基底膜对缓慢线性啁啾信号（2–10 千赫或 5–30 千赫）的响应，并使用滑动高斯窗分析响应而获得的。图 3B 并列展示了特征频率相近的沙鼠模型和人类模型的响应（沙鼠：4 千赫，人类：5.5 千赫）。结果显示，人类模型具有更尖锐的频率调谐，同时在不同声级下拥有更宽的响应增益范围。在 V-1D 模型中，这两种特性彼此耦合，因此，降低频率选择性会导致可实现的压缩程度减小。

图3C在较高特征频率下证实了这些观察结果，并展示了与已发表的基底膜振动数据的比较。图3C和3D重现了使用光学相干断层扫描（OCT）测得的活体沙鼠耳蜗基底圈频率响应（He 等人（2022））。实验数据和仿真的最佳频率均为19.5千赫；为简便起见，全文均将特征频率记作CF=20 千赫。图3C中与V-1D模型仿真的叠加结果证实，频率调谐相似，但模型在最佳频率处的压缩程度较弱，约低10分贝。采用调谐更尖锐的替代模型版本可以恢复完整的压缩范围（图3D），但代价是在低声级下产生不切实际的高频率选择性。增益与频率选择性之间的权衡可能使问题从压缩不足转变为调谐不当，但V-1D模型无法同时解决这两个方面。

III.2 V 模型中的短波增益补偿

$V^\star$ 模型通过添加一个作用于基底膜导纳之上的递归全极点滤波器（记为 β），实现部分增益补偿。图 4 展示了主动过程强度不同时， β 的幅值和相位。与预期一致， β 对长波区域对应的低频几乎没有影响（增益为 0 分贝），但在短波频率（例如接近特征频率处）表现出更显著的变化。该模型仅实现了通过两次连续回归得到的滤波器版本。第二次回归（径向基函数神经网络）的结果在幅值上可能与初始目标（S-2D 模型）相差 25%。如果相对于参考配置表示增益增强，回归会低估 β；若表示增益降低（即 $\beta < 1$），则会高估 β。这表明该方法仅适用于对基底膜导纳进行相对较小的修正。尽管如此， β 随频率和 G 值变化的总体趋势仍得到了良好捕捉。通过数值计算滤波器的极点并在 z 域中表示，对所得滤波器的稳定性进行了检查。未观察到任何不稳定情况。岭回归产生的滤波器系数较小，有助于提高滤波器的稳定性。

我们模拟了该模型在特征频率为 $= 20$ 千赫的位置对啁啾信号的响应，方法与此前对 V-1D模型所用的方法相同。结果如图 5 所示。导纳校正并未改变频率响应的尖锐程度，但将压缩范围扩大了 5 分贝。这一结果几乎完全实现了所需的峰值压缩（表现为随着声级升高，最大响应向较低频率移动）。对特征频率纯音（20 千赫）的响应压缩仍略显不足，差了 3 分贝。此外，模型实现的压缩仍伴随着耳蜗滤波器带宽的增宽，而实验数据中的品质因数仅降低了一半（20 分贝时为 $Q_{10dB} = 3.4$ ，10 分贝时为 $Q_{10dB} = 1.7$ ）。因此，与实验观测相比，高声级下的频率响应明显更宽。

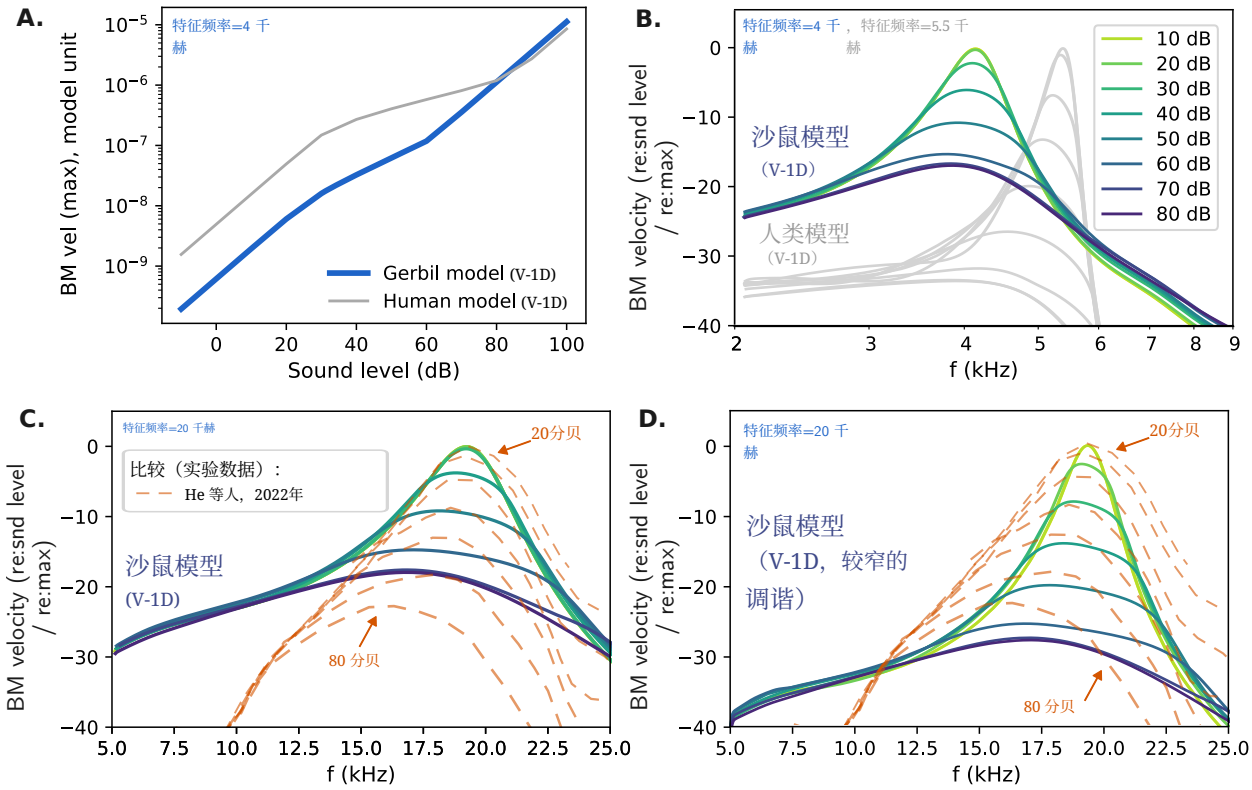


图3：**初始V-1D模型的特征响应**。**A.** 在特征频率为=4 千赫的位置，基底膜速度增长函数随刺激声级的变化（刺激为特征频率纯音）。**B至D**。 通过不同声级下呈现的缓慢线性啁啾信号所引起的基底膜速度响应，估算得到的V-1D模型频率响应（幅值）。**B** 沙鼠模型在特征频率为= 4 千赫位置的频率响应（彩色）。同时显示了人类模型在特征频率为=5.5千赫位置的响应（灰色）。**C.** 特征频率为=20 千赫位置的频率响应。声级的颜色编码与B中相同。图中叠加了沙鼠的光学相干断层扫描测量振动数据（He 等人（2022）），以虚线表示。对于实验数据，声级单位为声压级分贝，速度是相对于镫骨运动的速度。**D.** 与C相同，但采用了调谐更尖锐的模型版本。图B至D：‘re:snd level / re:max’表示频率响应已按照声级以及每个曲线组的最大峰值幅度进行归一化。

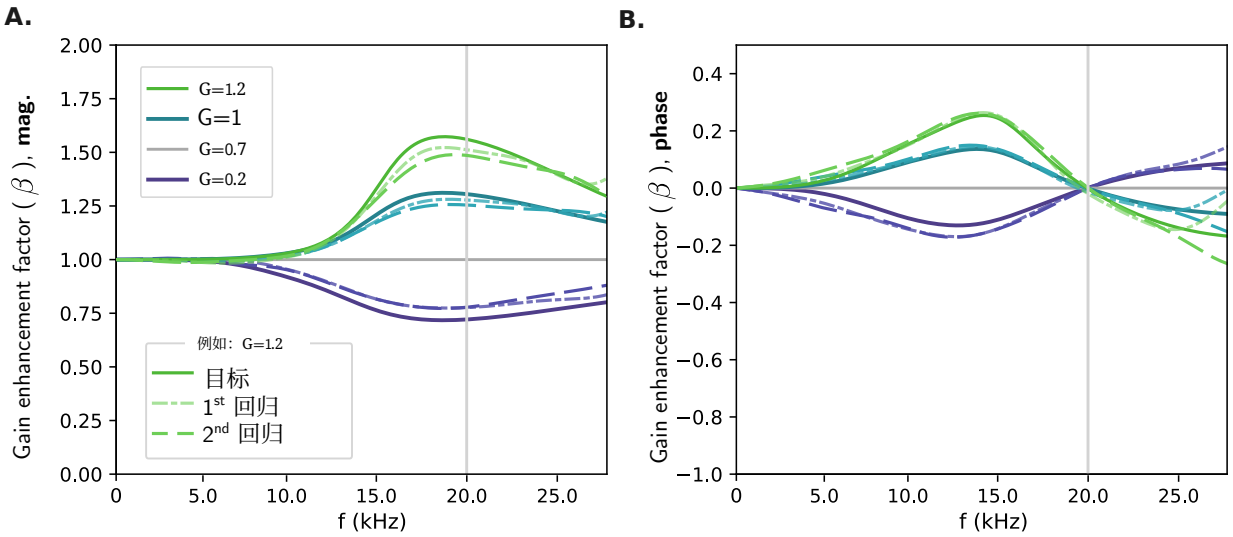


图 4：增益增强因子 β （**A.** 增益和 **B.** 相位），对应 G （活跃过程的强度）的四个不同取值。目标使用 Sisto 等针对 S-2D 模型所述的递归过程以及关系式 $\beta = \alpha/\alpha_0$ 计算，其中 α_0 为 $G_{REF} = 0.7$ （参考值）的压力聚焦因子。虚线曲线对应于用于将 β 近似为递归全极点滤波器的回归。点划线：第一次回归（岭回归）。虚线：第二次回归（径向基函数神经网络）。在两个面板中，竖线对应于基底膜特征频率 $\omega_{BM}/(2\pi)$。

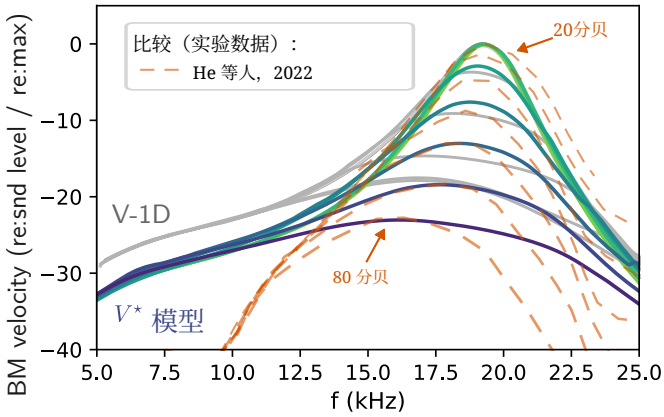


图 5： $V^\star$ 模型（特征频率=20千赫）对纯音响应的基底膜速度幅值，已按声级归一化。声级从 20 分贝到 80 分贝，步长为 10 分贝（颜色编码与图 3 相同）。V-1D模型的响应（灰线）和真实振动数据（虚线）复现自图 3。

III.3 模型校准：更新速率与压缩性增长

最后，我们展示另外两幅与 $V^\star$ 模型校准相关的图。这些图涉及模型实现的两个关键方面：增益补偿的时间更新，以及压缩非线性的校准。

$V^\star$ 中的因子 β 会随时间更新，其依据是输出强度 $\langle I_n \rangle[t]$的短时平均值，该值已在方法 II.4.3 中定义。为展示这一时间变化过程，图 6A 显示了特征频率为=20 千赫时对咔嗒声的响应。中间面板还显示了与活跃过程强度相对应的中间变量 G 随时间的变化。 G 大致跟随基底膜响应的包络（上下翻转），因为活跃过程在低强度时达到最大强度，而在高强度时则处于非活跃状态。图 6A 对应于模型所采用的更精细的 G 时间采样（$\Delta t = 5$微秒或 $f_s = 200$ 千赫）。这种采样使仿真计算量很大，因为每个时间步长都要重新计算 β 滤波器系数。与 V-1D 模型相比，采用这种采样时，计算时间增加了 20 倍。为解决这一限制，我们测试了其他版本的 $V^\star$ 模型，采用较慢的 β 滤波器更新速率。相应的脉冲响应见图 6B 和 6C。与较快更新速率（面板 A；基线）相比，时域响应上的差异不易察觉，但在右侧所示的频率响应中可以看出。当更新速率为 0.06 毫秒（比最快更新慢 12 倍）时，幅值与基线的差异最多达 ~3 分贝，我们认为差异过大。当更新速率为 0.03 毫秒（慢 6 倍）时，偏差不超过 ~1 分贝，我们认为这是计算效率与响应一致性之间的良好折衷。因此，我们将 0.03 毫秒的更新速率设为模型的默认值。与最快更新速率相比，这几乎将计算时间缩短了 6 倍，但仍约为 V-1D 模型计算时间的 5 倍。不过，只要将滤波器系数存储在查找表中——该表在首次运行时只需几秒钟即可计算完成——仿真时间便可与 V-1D 模型相当。上述计算时间不包括数据保存操作；在实际应用中，数据保存操作是时域 TL 模型的主要瓶颈，对 V-1D 和 $V^\star$ 模型均有影响。

方法部分末尾解释了压缩非线性的校准，图7A以沙鼠模型为例进行了说明。该过程依据的是在对应特征频率的位置对 20 kHz 纯音的响应。响应在第一个转折点之前呈线性，在两个转折点之间呈压缩性，而在高声级下再次呈线性。对于 V-1D模型，压缩区间只是两个线性区间之间的过渡，对应于系统从低阻尼转变为高阻尼。对于 $V^\star$ 模型，另一个压缩因子是压力聚焦效应：在第一个转折点处，该效应表现为增益增强，但在压缩区间内逐渐转变为增益降低。图7展示了两种模型的响应增长：A面板展示了校准过程所预期的增长函数，B面板则展示了通过仿真获得的实际响应。 $V^\star$ 模型将压缩区间延长了 10 分贝，这与实验数据更加吻合。

IV 讨论

IV.1 通过 $V^\star$ 模型改进压缩非线性

我们实现了一种方法，以扩大 Verhulst 等人（V-1D）时域模型的响应增益范围，从而得到改进后的 $V^\star$ 模型。该校正基于自回归滤波器和回归方法，在对应于非线性行波短波区域的频率区域中引入了额外的增益变化来源。 $V^\star$ 模型保留了 V-1D 模型的大部分特性，同时实现了频率选择性与增益之间的一定程度的解耦。我们通过沙鼠模型中对特征频率为 $=20$ 千赫处基底膜速度响应的仿真展示了这一点。我们证明，峰值响应压缩增加了 5 分贝（图5），发生压缩的声级范围扩大了 10 分贝（图7）。在开发 V-1D 模型期间，人们已指出耳蜗响应增益范围有限可能是一个问题；该模型最初是基于人类听觉设计的（Verhulst 等人（2018））。近期将模型改用于小型啮齿动物，尤其是沙鼠和小鼠时，这一问题显得更加突出。所提出的导纳校正

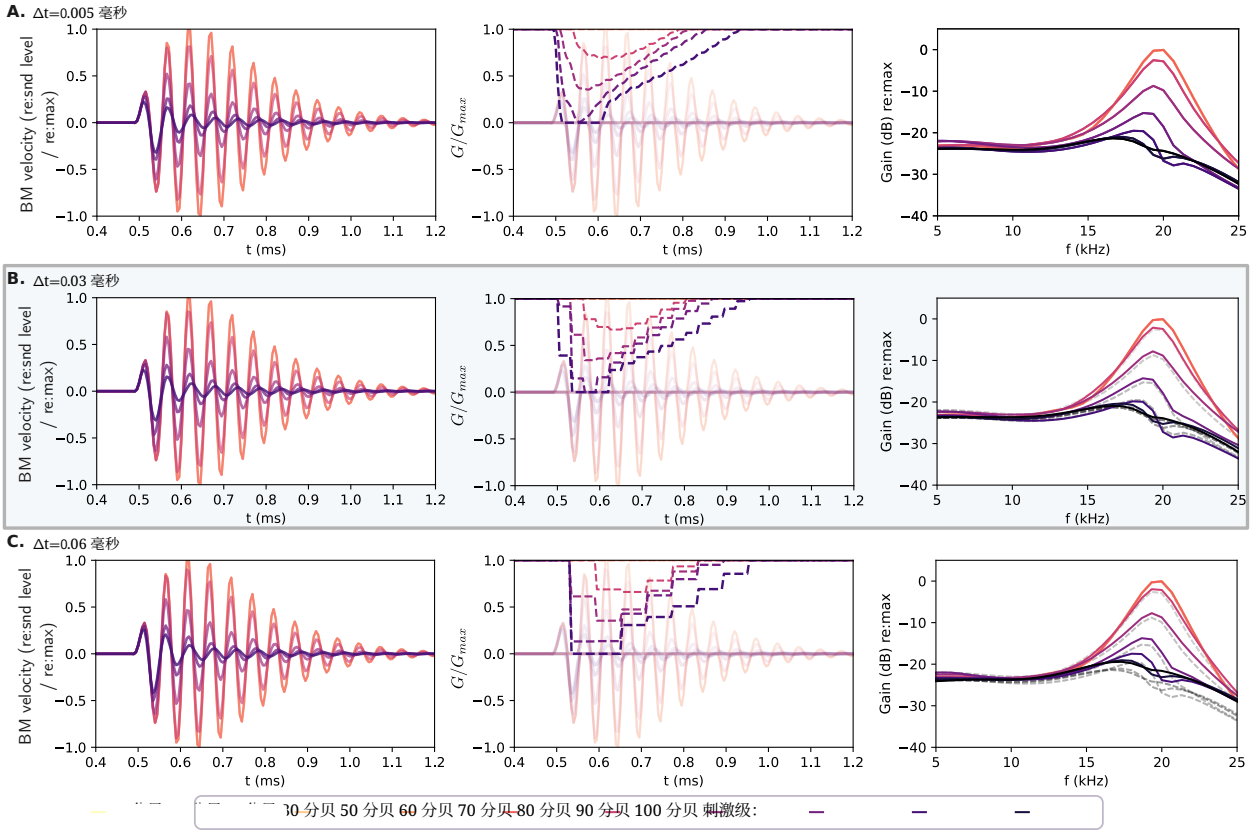


图6：**针对 $V\star$ 模型在不同更新速率下的短声响应**。**左：**特征频率=20 千赫位置处时域中的基底膜速度响应，已相对于刺激声级进行归一化。响应显示在 [-1, 1] 的标尺上，其中1对应于最敏感状态下的最大基底膜响应。咔嗒声在0.5毫秒时呈现。**中：** G/G_{max} 的时间进程（虚线曲线），表示用于计算增益补偿因子的活跃过程的强度。在同一图中，已绘出短声响应的轨迹（透明曲线）。**右：**与左图相对应的频率响应。上方面板中的轨迹以灰色虚线重现于下方各面板中，以供比较。每一行代表增益校正更新的不同时间分辨率：**A.** 0.005毫秒，**B.** 0.03毫秒**C.** 0.06毫秒。加框的一行（$\Delta t = 0.03$ 毫秒）对应于结果部分其他图中所使用的版本。

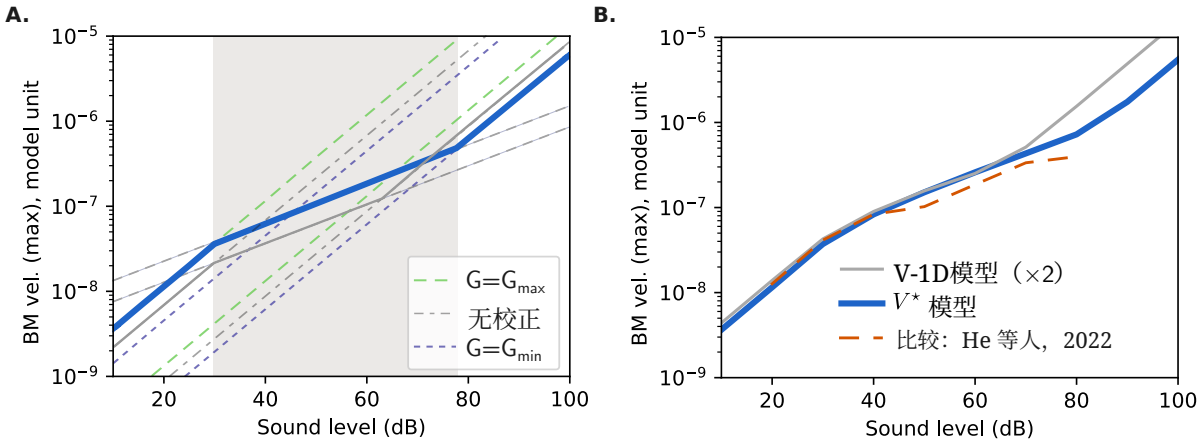


图7：**基底膜压缩性增长的校准**。**A.** 使用 $V^\star$ 模型的不同线性版本（虚线）模拟的基底膜速度增长函数。耳蜗分段的特征频率为20 kHz，刺激探针为20 kHz 纯音。响应分为两组：“起始极点”对应尖锐调谐模型，“饱和极点”对应适用于高声级的宽泛调谐模型。绿色曲线表示采用最大增益增强（$G = G_{max} = 1.3$）得到的结果，蓝色虚线表示采用最大增益降低（$G = G_{min} = 0$）得到的结果。此外，我们还以粗蓝实线表示非线性 $V^\star$ 模型的理论增长函数，以灰色实线表示非线性 V-1D 模型（无增益校正）的理论增长函数。灰色区域突出显示了压缩区间。**B.** 使用非线性 $V^\star$ 和 V-1D 模型得到的实际增长函数。为便于比较，V-1D模型的响应在垂直方向上进行了平移（乘以因子2）。图中还显示了实验数据（He 等人（2022））的增长函数（虚线）。

本研究为频率选择性较差的物种模型提供了一种实用方案，可补偿其压缩不足；如有需要，也可用于扩展调谐更尖锐物种的压缩范围。

然而，所提出的方法在实现现实的响应增益方面也存在一些局限。从图5中的峰尾比可以看出，模拟滤波器的整体增益远低于基底膜响应实验观测到的增益。 $V^\star$ 模型额外增加的5分贝增益不足以显著缓解这一问题。虽然我们的方法能够捕捉峰值增益随声级变化而产生的变化，但基线增益——定义为被动状态下的耳蜗响应增益——受时域实现的固有限制。这一点将在下一小节中进一步讨论。正因如此，我们将 β 因子用作自回归滤波器的目标值，其最大幅值介于0.5和2之间（图4），而没有采用压力聚焦因子 α；后者可在波峰处贡献约20分贝的增益（Sisto 等人，2021）。 β 因子反映的是与主动过程强度变化相关的 α 因子变化，而非绝对增益。因此， $V^\star$ 模型主要适用于考察耳蜗响应随声级变化的非线性模型。即使只考虑增益的相对差异，我们也发现，回归和滤波器设计步骤往往会减小 β 因子变化的幅度（图4）。

关于 $V^\star$ 模型的另一个问题是耳蜗基底与耳蜗顶端的行为差异。 β 滤波器的目标值是使用 S-2D 模型计算的，该模型假设耳蜗高度恒定。实际上，耳蜗的几何结构呈锥形，因此耳蜗顶端的压力聚焦因子（式 4）弱于基底（Altoè 和 Shera，2020）。因此，虽然在实践中，我们的方法适用于低于 20 千赫的特征频率（例如，我们也分析了 4 千赫时的模型响应），但在耳蜗顶端，该方法与耳蜗物理机制的关联就更值得质疑。在文章核心部分，我们使用沙鼠耳蜗第一圈的振动数据（He 等人（2022年））验证了模型，因为该处的实验数据相对丰富。现有的耳蜗顶端测量数据（Meenderink等人，2022）表明，顶端响应也呈现出较大的压缩，而频率

选择性变化有限。目前尚不清楚，经过调整的 $V^{\star}$ 模型能否准确呈现耳蜗顶端的物理机制。

IV.2 在时域传输线模型中实现高增益的难点 l

前文提到了时域传输线模型在重现宽带基底膜响应时难以实现较高峰尾比的问题。要理解为什么这对 $V^{\star}$ 模型仍然是个挑战，可以考察一维前行波的 WKB近似形式： $v_{BM}(x,\omega) \propto Y_{BM}^{3/4}(x,\omega)\exp(-j\int_0^x \kappa(x',\omega)dx')$ （Reichenbach 和 Hudspeth，2014年） [2]。v_{BM}中的增益可能直接来自导纳 Y_{BM}，也可能来自指数项。本文此前已多次提到基于导纳的增益所受到的限制：它受制于频率选择性与增益之间的权衡，这一现象在典型带通滤波器中广为人知。这里，我们转而关注如何通过第二个因子——取决于复波长空间积分的指数项——来实现增益。

在敏感耳蜗中，脉动频率为 ω 的单色波会在其峰值基底侧的一小段区域内得到放大，此处复波数 $\kappa(x',\omega)$ 的相位为正。假设尺度对称性成立，这也意味着在任意给定的耳蜗位置 x，在低于基底膜特征频率的一个狭窄频率范围内， $\kappa(x,\cdot)$ 的相位为正（Altoè等人，2021年）。因此，若将 $\kappa(x,\cdot)$ 视为滤波器，它必然包含一个相位导数为正的区域，对应于负群延迟。这一特性也适用于基底膜导纳。然而，因果滤波器难以实现负群延迟；一旦出现负群延迟，也是源于两个必然具有正延迟的波包之间的相位干涉。实际上，只有当某个频率区域的负群延迟由相邻频率区域中幅值更大、群延迟为正的部分所补偿时，才可能出现负群延迟区域。例如，图4中用于 $G>1$ 的 β 滤波器就呈现了这一模式：在 14 千赫以下，群延迟为负（相位导数为正）；而在 14 千赫以上，则出现群延迟为正（相位导数为负）的区域，且该区域的响应幅值更大。本段所述的限制还因以下事实而更加突出：具有较大峰尾比的因果滤波器，其峰值区域通常对应较大的正群延迟。因此，基底膜速度增益的两个来源——基底膜导纳，或负电阻区域引起的波传播——无法在同一频带内叠加。

因此，要设计出既具有较高峰尾比、又仅在峰值区域产生适度正相移的因果滤波器颇具挑战。这严重限制了传输线模型的时域实现，因为它完全依赖因果分量。然而，这就引出了一个问题：对于真实耳蜗，这一论点又有何不同？这个问题值得探讨，因为物理系统同样受到因果性约束（即克拉默斯-克罗尼希关系（Shera 2007）），但 v_{BM} 数据却表现出较高的峰尾比。事实上，对于耳蜗行波而言，因果性约束并不像乍看之下那么严格，因为除了局部“点因果性”之外，还应考虑全局波因果性，而前者比后者更严格。局部“点因果性”指局部变量之间的因果关系；全局波因果性则还考虑波效应，这些效应能够以高达波前速度的速度传播，因此对局部观察者而言近乎瞬时。对于时域实现，我们将这些因子α 或 β 视为更新后的基底膜导纳的一部分。然而，从更准确的物理角度来看，这些因子应与局部基底膜导纳区分开来。具体而言，压力聚焦因子 α 与前向压力波相关；前向压力波在沿耳蜗传播时会将不同频率分量分开。波的色散特性意味着，每个波分量受到的压力聚焦影响各不相同，且独立于局部基底膜机械响应所引入的延迟。因此，将 α 表示为滤波器时，它不必满足因果性，即不必满足局部“点因果性”，但必须符合全局波因果性。不受这一约束后，它便能以近乎瞬时的方式显著增强与短波分量相关的频率响应，在群延迟极短的情况下实现高达 $+20$ 分贝的增益。这一区别虽微妙却至关重要，凸显了将二维频域模型改造为一维时域模型所面临的挑战。这一复杂性无法从方程中直接看出，而是源于对因果性的非平凡解读。

[2]该公式适用于流体阻抗 $Z_f = \frac{2j\omega\rho}{H}$ 被视为常数的箱体模型。对于长波近似下的渐缩箱体模型，如Zweig 1991（因而亦如V-1D模型）中的情形，归一化因子还包括$1/\sqrt{H}$；此处将其忽略，因为它取决于位置，而不取决于 ω。

即便是基底膜导纳 Y_{BM} 的局部因果性假设，也值得商榷，至少当它在行波框架内描述基底膜机械响应，而非将基底膜视为孤立片段时，情况如此。S-2D 模型很好地说明了这一点：其中 Y_{BM} 包含一个黏性阻尼项，该项取决于波形，并受到波的色散特性影响。更具体地说，导纳中的黏性应力与压力场的二阶空间导数有关。用 v_{BM} 表示时，它又会引入因子 α （式6），该因子在短波区域与 k 成正比。其他模型中也存在波长相关导纳：例如，Elliott 等人的近期研究通过盖膜的黏弹性模型纳入了纵向耦合（Elliott 等人，2024）。放宽局部点因果性约束的导纳模型，可以消除时域实现中观察到的导纳相位振荡（特征频率以下）（Shera，2001），从而使负电阻区域在频域和空间域中得以扩展。

由于难以用因果稳定模型模拟负电阻区域，时域实现似乎有必要利用基底膜导纳的共振来产生高响应峰。对于时域模型，人们普遍预期，对特征频率纯音的响应波会在耳蜗位置达到波峰，此处的特征频率等于 CF；也就是说，导纳中的刚度项和质量项相互抵消。正如方法部分 II.3 中所述，这一点对于 V-1D 模型尤其如此。尽管这一特性看起来顺理成章，甚至近乎同义反复，但它并非二维频域模型必然具备的特征（S-2D 模型就不具备）。如果在特征位置基底侧的某个区域内波数急剧增加，那么增益（通过 α）和阻尼（通过黏性阻尼）的增加可能足以使耳蜗的刚度主导区域在波到达 CF 位置之前产生一个高峰，随后急剧滚降（Sisto 等人，2021；Deloche 等人，2025；Sisto 和 Moleti，2024）。行波峰出现在刚度主导区域，似乎更符合耳蜗内压力与基底膜位移的同步测量结果（Dong 和 Olson，2013），不过这一问题仍存在争议。

IV.3 V 模型中的两种非线性

$V^\star$ 模型通过两种不同的实现策略实现压缩。在本文最后的讨论中，我们将探讨这两种策略的差异，以及这种双重方法在其他情境中的潜在用途。

采用增益增强因子 β 的解与最初的基底膜导纳模型不同，因为它依赖数值方法，而非解析推导。最初的 V-1D 导纳模型只包含较少的参数（式 11），这些参数可直接通过刚度重量比、阻尼、慢反馈项的幅值或延迟来解释。除慢反馈项外，阻抗项涉及时间导数（最高至二阶），对应于（准）瞬时的局部动力学。相比之下，递归的 β 滤波器包含对之前 32 个历史值的加权求和，因此能够实现更慢的动力学。由此得到的滤波器频率响应更平滑、模式化程度更低，且可以更灵活地调整，以匹配所需的宽带形状。我们采用 S-2D 模型计算出的压力聚焦因子作为目标，但也可以采用其他模型得到的滤波器形状。重要的是，本文所述方法旨在调整导纳，而非复现其特征形式；在 V-1D 模型中，这种特征形式源自经过精心控制的共振。因此，V-1D 导纳方程仍然是 $V^\star$ 传输线模型的核心。这些方程受到主要极点轨迹的约束，以维持基底膜速度过零点近乎不变（Shera，2001）。一种可能增加特征频率处压缩的方法，是随着声级升高而降低基底膜导纳的局部共振频率。这样会加剧波峰向基底侧偏移，并使特征频率纯音的响应增益随声级升高而更大幅度地下降。然而，这种方法也会改变导纳相位响应，并破坏过零点的近乎不变性。在一维模型中纳入压力聚焦的另一种方法，是使用格林函数形式，对沿耳蜗分隔分布的压力源贡献进行积分（Diependaal 和 Viergever；Duifhuis，2012；Ver Hulst），但这种方法的计算成本非常高。

最后，这两种压缩非线性的实现方式有所不同，因为它们体现了不同类型的非线性动力学。与 V-1D模型类似， $V^\star$ 模型主要由瞬时非线性驱动；这类非线性被认为能够以逐周期的方式增强耳蜗运动（Dewey等人2021）。但增益增强

因子 β基于反馈回路，也会引入非瞬时非线性。压力聚焦的强度遵循基底膜速度包络的垂直翻转形式（图 6）。实际上，这种选择加快了计算速度，因为滤波器系数无需在每个时间步更新。但这种选择也有其物理依据：波数的变化不可能瞬时发生，必然需要在一个时间窗口内进行积分。尽管大多数耳蜗模型都包含瞬时非线性，但主动过程是否也可能遵循一种基于信号包络的反馈回路（类似于自动增益控制方案），一直存在争论（Van Der Heijden 2005；Cooper和van der Heijden 2016；Altoè等人2017）。在我们的模型中，部分非线性由反馈回路驱动，但这些非线性最初源于伴随主动过程变化而出现的一种被动现象（压力聚焦）。因此，对于 $V^\star$ 模型，非线性压缩响应背后的机制仍被认为是瞬时的。无论其源自被动过程还是主动过程，耳蜗中非线性动力学的类型都具有重要意义。例如，瞬时非线性或基于包络的控制回路会产生效应大小（Duifhuis，2012）或相位特性（Van Der Heijden 2005）不同的失真产物。在这方面，$V^\star$ 模型或其变体能够模拟复杂的混合动力学，因此可能有助于研究耳蜗力学中更深入的时间特性。

总之，我们提出了一种方法，可在一维传输线模型中实现增益响应的额外变化因子。该因子部分再现了二维模型在短波区域观察到的响应“增强”。由此得到的 $V^\star$ 模型在波峰处产生了高出 5 分贝的增益（或压缩）。事实证明，该方法有助于开发针对小型哺乳动物、尤其是啮齿动物生理特征定制的传输线模型。作为转化研究方法的一部分，开发此类模型很有价值，因为常见的实验动物能为正常听力和听力受损提供大量研究见解。不过，要实现这种校正，需要采用涉及回归和滤波器设计方法的复杂方案。总体而言，这项工作凸显了将波长依赖现象转化为时域传输线实现、同时保持计算效率所面临的挑战。

致谢

本项目由 FWO 项目“Audimod”（G068621N）和 ERA-NET 项目“CoSySpeech”（G0H6420N）资助。

附录

A. 最小二乘回归代价函数的完整形式

第一次回归用于估计给定参数对 $[b_k]_{k=1\cdots32}$的系数 (ω_{BM}, G)。在此步骤中，我们对频率 $(0, f_s/(2m_2)\cdots, m_1 f_s/(2m_2), \cdots, (m_2-1)f_s/(2m_2))$进行均匀离散化，覆盖区间 $[0, f_s/2]$，其中 $f_s/2$ 是传输线仿真的奈奎斯特频率。在我们的案例中，f_s =为 200 千赫，且 $m_2 = 512$。离散频率集合进一步划分为 $(0, \cdots, (m_1-1)f_s/(2m_2))$和$(m_1 f_s/(2m_2), \cdots, (m_2-1)f_s/(2m_2))$，分别对应低于和高于 $\omega_{cut} = 1.3\ \omega_{BM}$的频率。

我们定义以下“特征”矩阵：

$$X_{l,re} = [\cos(\pi(km)/m_2)]_{m=0\cdots m_1-1,\, k=1\cdots32}\,, \tag{19a}$$

$$X_{l,im} = -[\sin(\pi(km)/m_2)]_{m=0\cdots m_1-1,\, k=1\cdots32}\,, \tag{19b}$$

$$X_{r,re} = [\cos(\pi(km)/m_2)]_{m=m_1\cdots m_2-1,\, k=1\cdots32}\,, \tag{19c}$$

$$X_{r,im} = -[\sin(\pi(km)/m_2)]_{m=m_1\cdots m_2-1,\, k=1\cdots32}\,. \tag{19d}$$

l 表示“左”，r 表示“右”。根据公式 13，我们有

$$Y_{l,re} = \mathrm{Re}(1 - 1/\beta) \approx 1 - \mathrm{Re}(1/\hat{\beta}) = X_{l,re} b, \tag{20a}$$
$$Y_{l,im} = \mathrm{Im}(1 - 1/\beta) \approx -\mathrm{Im}(1/\hat{\beta}) = X_{l,im} b, \tag{20b}$$

其中， b 是系数 (b_k)的向量。

完整的LSE代价函数如下:

$$LSE(b) = \left\| Y_{l,re} - X_{l,re} b \right\|_2^2 + \left\| Y_{l,im} - X_{l,im} b \right\|_2^2 + \lambda_1 \left\| b \right\|_2^2 + \lambda_2 \left\| \max(0, \frac{X_{r,Im} b}{1 - X_{r,Re} b}) \right\|_2^2 \tag{21}$$

其中， $\|\cdot\|_2$ 是欧几里得范数。该代价函数包含两个惩罚项，分别由 λ_1 和 λ_2加权，这些惩罚项^防止 b_k 中的系数过大，以及频率高于 ω_{cut}时 β 中的正相位值过大。在第二个惩罚项中， $\frac{X_{Im} b}{1 - X_{Re} b}$ 对应于 $\mathrm{Im}(\hat{\beta})/\mathrm{Re}(\hat{\beta})$，使用特征矩阵表示，并采用逐元素除法。

使用SciPy的函数 `leastsq` 最小化了LSE代价函数（公式21），该函数基于Levenberg–Marquardt（LM）算法。

B. 强制校正因子在低频下的增益为零

^上述方法（即使用公式21）存在一个问题：当 $\omega \to 0$时，无法保证估计的滤波器 $\beta(\omega)$ 的幅值为1。一种强制 $\beta(0) = \left(1 - \sum_{k=1}^{32} b_k\right)^{-1} = 1$ 的方法是将系数 $[b_k]$ 约束在超平面 $\sum_{k=1}^{32} b_k = 0$上，但我们发现这种解决方案过于僵化，导致了不理想的行为。

相反，我们在用于确定滤波器系数的回归中加入了增益因子 $K = 1 + \epsilon$，并在优化后对其进行调整，以确保频率 0 处的增益为零（以分贝计）。具体而言，式 13 替换为:

$$\beta(z^{-1}) \approx \frac{1 + \epsilon}{1 - \sum_{k=1}^{32} b_k z^{-k}}. \tag{22}$$

由此，式 20 更新为:

$$\mathrm{Re}(1 - 1/\beta) - \epsilon \mathrm{Re}(1/\beta) \approx X_{l,re} b, \tag{23a}$$
$$\mathrm{Im}(1 - 1/\beta) - \epsilon \mathrm{Im}(1/\beta) \approx X_{l,im} b, \tag{23b}$$

这将得到修改后的 LSE 函数:

$$LSE(b, \epsilon) = \left\| Y_{l,re} - X_{l,re} b + (Y_{l,re} - 1)\epsilon \right\|_2^2 + \left\| Y_{l,im} - X_{l,im} b + Y_{l,im} \epsilon \right\|_2^2 + \lambda_1 (\left\| b \right\|_2^2 + \epsilon^2) + \lambda_2 \left\| \max(0, \frac{X_{r,Im} b}{1 - X_{r,Re} b}) \right\|_2^2 . \tag{24}$$

扩展的LSE代价如前所述使用LM算法最小化，除 $[b_k]$ 系数外，还将 ϵ 作为回归变量。优化完成后，我们按照 $\epsilon \leftarrow -\sum_{k=1}^{32} b_k$更新 ϵ，以设定传递函数$K = 1 + \epsilon$ 的期望增益。

参考文献

Alessandro Altoè与Christopher A. Shera。耳蜗耳角：听觉信号处理中音调定位变化的几何起源。*Scientific Reports*，10：20528，2020年11月。ISSN 2045-2322。doi：10.103 8/s41598-020-77042-w。Alessandro Altoè、Ville Pulkki与Sarah Verhulst。传输线耳蜗模型：精度与效率的改进。美国声学学会期刊，136（4）：EL302–EL308，2014年9月。ISSN 0001-4966。doi：10.1121/1.4896416。Alessandro Altoè、Karolina K. Charaziak与Christopher A. Shera。耳蜗非线性动力学：自动增益控制还是瞬时阻尼？美国声学学会期刊，142（6）：3510–3519，2017年12月。ISSN 0001-4966，1520-8524。doi：10.1121/1.5014039。Alessandro Altoè、Karolina K. Charaziak、James B. Dewey、Arturo Moleti、Renata Sisto、John S. Oghalai与Christopher A. Shera。难以捉摸的耳蜗滤波器：耳蜗跨频掩蔽的波动起源。*Journal of the Association for Research in Otolaryngology 2021*年，第1–18页，2021年10月。ISSN 1438-7573。doi：10.1007/S10162-021-00814-2。Deepak Baby、Arthur Van Den Broucke与Sarah Verhulst。面向实时应用的人类耳蜗力学与滤波器调谐卷积神经网络模型。《自然·机器智能》，3（2）：134–143，2021年2月。ISSN 2522-5839。doi：10.1038/s42256-020-00286-8。Karolina K. Charaziak与Christopher A. Shera。由短声和频率扫掠诱发的反射源发射：不同声级的比较。*JARO: Journal of the Association for Research in Otolaryngology*，22（6）：641–658，2021年12月。ISSN 1525-3961。doi：10.1007/s10162-021-00813-3。Nigel P. Cooper与Marcel van der Heijden。耳蜗非线性动力学。载于Pim van Dijk、Deniz Baskent、Etienne Gaudrain、Emile de Kleine、Andreas Wagner与Casper Lanting编，《正常与受损听力的生理学、心理声学与认知》，第267–273页。施普林格，柏林，2016年。François Deloche、Morgan Thienpont、Arturo Moleti、Renata Sisto与Sarah Verhulst。盖膜横向黏弹性阻尼的主动控制：行波放大的第二种机制？*Hearing Research*，464：109320，2025年。ISSN 03785955。doi：10.1016/j.heares.2025.109320。James B. Dewey、Alessandro Altoè、Christopher A. Shera、Brian E. Applegate与John S. Oghalai。耳蜗外毛细胞电动性在体内高频条件下逐周期增强柯蒂氏器运动。《美国国家科学院院刊》，118（43）：e2025206118，2021年10月。doi：10.1073/pnas.2025206118。Rob J. Diependaal与Max A. Viergever。非线性主动二维耳蜗模型：时域解。85（2）：803–812。ISSN 0001-4966。doi：10.1121/1.397553。URL https: // doi.org/10.1121/1.397553。Rob J. Diependaal、H. Duifhuis、H. W. Hoogstraten与Max A. Viergever。求解时域一维耳蜗模型的数值方法。美国声学学会期刊，82（5）：1655–1666，1987年11月。ISSN 0001-4966，1520-8524。doi：10.1121/1.395157。Wei Dong与Elizabeth S. Olson。耳蜗放大及其激活的检测。《生物物理学杂志》，105（4）：1067–1078，2013年8月。ISSN 0006-3495。doi：10.1016/j.bpj.2013.06.049。Hendrikus Duifhuis。耳蜗力学：非线性耳蜗时域分析导论。施普林格美国，马萨诸塞州波士顿，2012年。ISBN 978-1-4419-6116-7 978-1-4419-6117-4。doi：10.1007/978-1-4419-6117-4。Stephen Elliott、Riccardo Marrocchio与Karl Grosh。柯蒂氏器中纵向耦合的形式。《*AIP*会议论文集》，3062（1）：020014，2024年2月。ISSN 0094-243X。doi：10.1063/5.0189306。Elika Fallah、C. Elliott Strimbu与Elizabeth S. Olson。耳蜗内运动与局部耳蜗微音器电位的非线性：豚鼠与沙鼠的比较。*Hearing Research*，405：108234，2021年6月。ISSN 03785955。doi：10.1016/j.heares.2021.108234。

Wenxuan He、George Burwood、Edward V. Porsov、Anders Fridberger、Alfred L. Nuttall 和 Tianying Ren。活体沙鼠耳蜗基底圈中网状板和基底膜在横向的振动。*Scientific Reports*，12(1):19810，2022年11月。ISSN 2045-2322。doi: 10.1038/s41598-022-24394-0。Sebastiaan W. F. Meenderink、Xiaohui Lin、B. Hyle Park 和 Wei Dong。声音诱发的振动使沙鼠耳蜗低频顶端区域的柯蒂氏器复合体发生形变，从而实现正常听力：声音诱发的振动使柯蒂氏器复合体发生形变。*Journal of the Association for Research inOtolaryngology: JARO*，23(5):579–591，2022年10月。ISSN 1438-7573。doi: 10.1007/s10162-022-00856-0。Marcus Müller。成年及发育中的蒙古沙鼠的耳蜗位置-频率图。*HearingResearch*，94(1):148–156，1996年5月。ISSN 0378-5955。doi: 10.1016/0378-5955(95)00230-8。T. Reichenbach 和 A. J. Hudspeth。听觉物理学：内耳的流体力学和主动过程。*Reports on Progress in Physics*，77(7):076601，2014年7月。ISSN 0034-4885，1361-6633。doi: 10.1088/0034-4885/77/7/076601。Mario A. Ruggero、Nola C. Rich、Alberto Recio、S. Shyamla Narayan 和 Luis Robles。毛丝鼠耳蜗基底部对纯音的基底膜响应。美国声学学会期刊，101(4):2151–2163，1997年4月。ISSN 0001-4966。doi: 10.1121/1.418265。Amin Saremi、Rainer Beutelmann、Mathias Dietz、Go Ashida、Jutta Kretzberg 和 Sarah Verhulst。七种人类耳蜗滤波器模型的比较研究。美国声学学会期刊，140(3):1618–1634，2016年9月。ISSN 0001-4966。doi: 10.1121/1.4960486。Christopher A. Shera。基底膜点击响应中精细时间结构的强度不变性：对耳蜗力学的启示。美国声学学会期刊，110(1):332–348，2001年7月。ISSN 0001-4966。doi: 10.1121/1.1378349。Christopher A. Shera。带有转折的激光放大：来自整个耳蜗的行波传播和增益函数。美国声学学会期刊，122(5):2738–2758，2007年11月。ISSN 0001-4966，1520-8524。doi: 10.1121/1.2783205。Christopher A Shera 和 Karolina K Charaziak。耳蜗频率调谐和耳声发射。*ColdSpring Harbor Perspectives in Medicine*，9(2)，2019年。ISSN 21571422。doi: 10.1101/cshperspect.a033498。Christopher A. Shera 和 George Zweig。一种对称性抑制了耳蜗灾变。美国声学学会期刊，89(3):1276–1289，1991年3月。ISSN 0001-4966，1520-8524。doi: 10.1121/1.400650。Christopher A Shera、John J Guinan 和 Andrew J Oxenham。根据耳声发射和行为测量重新估计人类耳蜗调谐。99(5):3318–3323。ISSN 0027-8424。doi: 10.1073/pnas.032 675099。URL http://www.ncbi.nlm.nih.gov/pubmed/11867706。Renata Sisto 和 Arturo Moleti。耳蜗音调定位之谜：具有“非谐振”响应峰的谐振传输线。*JASA Express Letters*，4(7):074401，2024年。ISSN 2691-1191。doi: 10.1121/10.002 8020。Renata Sisto、Daniele Belardinelli 和 Arturo Moleti。流体聚焦和黏度使耳蜗响应具有高增益和稳定性。美国声学学会期刊，150(6):4283–4296，2021年12月。ISSN 0001-4966。doi: 10.1121/10.0008940。Renata Sisto、Daniele Belardinelli、Alessandro Altoè、Christopher A. Shera 和 Arturo Moleti。耳蜗响应理论建模中的关键三维黏性流体动力学贡献。美国声学学会期刊，153(1):77–86，2023年1月。ISSN 0001-4966。doi: 10.1121/10.0016809。Morgan Thienpont、Francois Deloche、Sarineh Keshishzadeh、Danill Kiselev、Jérôme Bourien、Jean-Luc Puel、Brad Buran、Naomi Bramhall 和 Sarah Verhulst。将人类听觉外周计算模型迁移至沙鼠和小鼠，用于比较听觉研究。载于*Mechanics of HearingWorkshop 2024 (MoH 24)*，安娜堡，美国，2024年6月。Zenodo。doi: 10.5281/zenodo.13769392。

Marcel Van Der Heijden。耳蜗增益控制。美国声学学会期刊，117(3): 1223–1233，2005年3月。ISSN 0001-4966，1520-8524。doi: 10.1121/1.1856375。Alejandro Osses Vecchi、Léo Varnet、Laurel H. Carney、Torsten Dau、Ian C. Bruce、Sarah Verhulst 和 Piotr Majdak。八种人类单耳处理听觉模型的比较研究。*Acta Acustica*，6: 17，2022年。ISSN 2681-4617。doi: 10.1051/aacus/2022008。Henri Ver Hulst。将临界振荡器的物理机制与耳蜗行波模型相结合。Sarah Verhulst、Torsten Dau 和 Christopher A. Shera。用于瞬态刺激和人类耳声发射的非线性时域耳蜗模型。美国声学学会期刊，132(6): 3842–3848，2012年12月。ISSN 0001-4966。doi: 10.1121/1.4763989。Sarah Verhulst、Hari M. Bharadwaj、Golbarg Mehraei、Christopher A. Shera 和 Barbara G. Shinn- Cunningham。对宽带刺激下人类听觉脑干反应的功能建模。美国声学学会期刊，138(3):1637–1659，2015年9月。ISSN 0001-4966。doi: 10.1121/1.4928305。Sarah Verhulst、Alessandro Altoè 和 Viacheslav Vasilkov。人类听觉外周的计算建模：听神经反应、诱发电位和听力损失。*Hearing Research*，360:55–75，2018年3月。ISSN 0378-5955。doi: 10.1016/j.heares.2017.12.018。Eric Verschooten、Christian Desloovere 和 Philip X. Joris。人类耳蜗具有高分辨率频率调谐，但不具备时间编码能力。*PLOS Biology*，16(10):e2005164，2018年10月。ISSN 1545-7885。doi: 10.1371/journal.pbio.2005164。G. Zweig。测定柯蒂氏器的阻抗。美国声学学会期刊，89(3):1229–1254，1991年3月。ISSN 0001-4966。doi: 10.1121/1.400653。