

# Auralization of Electric Vehicles for the Perceptual Evaluation of Acoustic Vehicle Alerting Systems

Leon Müller\*, Wolfgang Kropp

Division of Applied Acoustics, Chalmers University of Technology  
41296 Gothenburg, Sweden

\* Correspondence: leon.mueller@chalmers.se; Tel.: +46-707135619

## 1 Abstract

Electric vehicles (EVs) typically emit little noise at low driving speeds, which increases the risk of accidents for vulnerable road users such as pedestrians. To reduce this risk, regulations demand that newly sold EVs have to be equipped with an acoustic vehicle alerting system (AVAS), which radiates artificial warning sounds. Developing AVAS sounds that provide a sufficient warning capability while limiting traffic noise annoyance requires laboratory listening experiments; such experiments need accurate auralization methods. Even though several universal auralization tools are already established in the research field, those frameworks require additional data to simulate EVs. This paper presents an electric vehicle auralization toolchain combined with an open-access database, including AVAS measurements, synthesis algorithms, and numerically calculated sound source directivities for three different electric passenger cars. The auralization method was validated numerically and in a listening experiment, comparing simulated EV passages to binaural in-situ recordings. The results of this perceptual validation indicate that stimuli generated with the presented method are perceived as slightly less plausible than in-situ recordings and that they result in a similar distribution of annoyance ratings but a higher perceived vehicle velocity compared to the reference recordings.

## 29 1 Introduction

Electric vehicles (EVs) have become increasingly popular in the last few years, especially in urban environments [1]. EVs radiate significantly less sound at low driving speeds than internal combustion engine vehicles [2]. This reduced noise emission is, in general, beneficial as it is known that high road traffic noise levels can cause serious health problems, such as chronic annoyance, sleep disturbance, and cardiovascular diseases [3]. However, the low noise level of electric vehicles also increases the risk of accidents for pedestrians and vulnerable road users, such as the visually impaired, who can not localize approaching

EVs using acoustic cues [4]. To reduce this risk, recently implemented regulations demand that all newly produced electric vehicles have to be equipped with an acoustic vehicle alerting system (AVAS) [5], i.e., a loudspeaker that radiates artificial warning sounds indicating the vehicle's location and driving speed.

A challenge in designing these AVAS sounds is to achieve a sufficient warning effect while simultaneously limiting the negative impact on the acoustic environment. Therefore, a fundamental requirement for developing efficient AVAS systems is to understand which sound characteristics are relevant to localize a vehicle acoustically and how those sounds affect bystanders. A common approach to studying such psychoacoustic effects is to conduct listening experiments in controlled acoustic environments. Such studies require an accurate reproduction of the acoustic scenes of interest, a process that is also referred to as auralization [6].

The limited number of existing AVAS-related psychoacoustic studies often relied on auralization methods based on static recordings of EV passages obtained by, e.g., placing a microphone array on the side of a test track [7]. This approach generally results in realistic stimuli since it automatically includes complex acoustic properties, such as source directivity and sound propagation path. However, performing real-life recordings can be time- and resource-intensive and does not provide the flexibility of changing parameters such as vehicle speed or the AVAS signal after conducting the measurement. Other researchers combined recordings of microphones placed on a moving vehicle [8] and self-synthesized AVAS signals [9] with a general auralization model, resulting in more flexible simulations. Nevertheless, those auralizations do not necessarily include an accurate source directivity model, which can be expected to be perceptually relevant, especially for recently developed directional AVAS systems [10]. Even though several commercial and open-source auralization frameworks for rendering virtual acoustic scenes are already available [11–14], only few AVAS signals or EV directivity models are currently openly accessible to the scientific community, limiting the possibilities for researchers

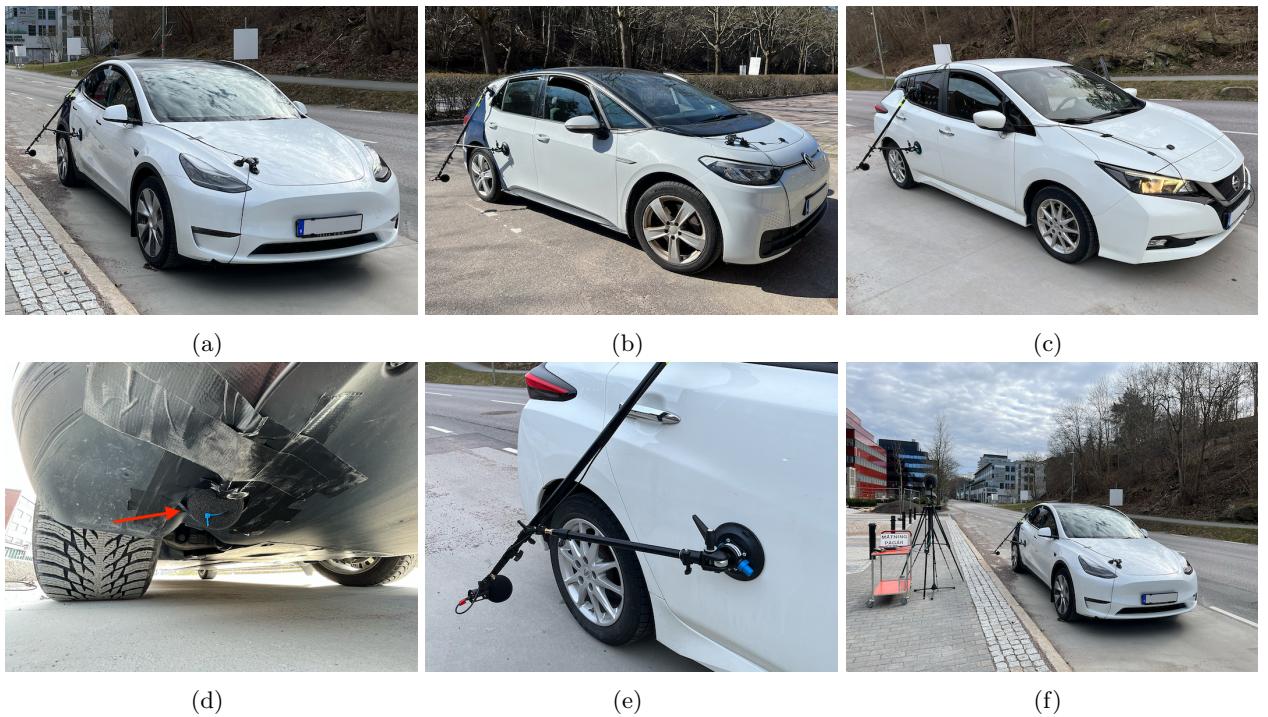


Figure 1: Vehicle A - Tesla Model Y 2021 (a), vehicle B - Volkswagen ID.3 Pro Performance 2021 (b), vehicle C - Nissan Leaf 2018 (c), microphone placed in front of vehicle A AVAS speaker (d), microphone placed in front of vehicle C tire (e), and measurement setup on roadside (f).

to conduct AVAS-related psychoacoustic studies. Additionally, most existing tire/road noise auralization methods are designed for velocities above 25 km/h, as combustion engine noise is typically considered dominating at lower speeds [15]. However, electric vehicles radiate significantly less motor noise than internal combustion engine vehicles; hence, even low-speed rolling noise can become audible and should, therefore, be included in the auralization.

This work offers resources and methods for researchers to conduct listening experiments related to Acoustic Vehicle Alerting Systems. It introduces an open-access database containing AVAS and tire-noise recordings, along with AVAS synthesis models for three electric vehicles. These models are paired with an auralization framework designed for headphone-based reproduction. The paper begins by outlining the measurement process and examining the characteristics of the recorded AVAS and tire/road noise signals in [Section 2](#). Building on these measurements, [Section 3](#) describes techniques for synthesizing AVAS and tire/road noise signals, including methodologies for modeling their radiation directivities and sound propagation. In [Section 4](#), these methods are applied to re-create measured vehicle passages, allowing for numerical validations through comparisons with the reference recordings. Finally, [Section 5](#) presents an assessment of the perceptual quality of the auralization results through a laboratory listening experiment.

## 2 Measurements

In order to obtain reference data for the auralization model, pass-by measurements of three different electric vehicles were conducted: a Tesla Model Y 2021 (vehicle A, [Figure 1a](#)), a Volkswagen ID.3 Pro Performance 2021 (vehicle B, [Figure 1b](#)) and a Nissan Leaf 2018 (vehicle C, [Figure 1c](#)). These vehicles can be classified as small to medium-sized electric passenger cars, which all utilize a single AVAS loudspeaker mounted in the front bumper. All vehicles were equipped with radial non-studded winter tires with an external rolling noise value of 72 dB according to EU regulation 2020/740 [16].

### 2.1 Methods

AVAS signals of each vehicle were measured by placing a microphone in front of the AVAS loudspeaker as shown in [Figure 1d](#). Additionally, a second microphone was mounted at 40 cm distance perpendicular to the tire as shown in [Figure 1e](#) to record isolated tire/road noise. The sound pressure at a stationary observer position was measured by placing a HEAD acoustics HMS V artificial head and a microphone at the roadside, as shown in [Figure 1f](#). All microphones were free-field equalized and of the type GRAS 46AE with 90 mm foam windscreens. Vehicle velocity and position relative to the roadside observer were recorded via GPS, allowing for the exact repro-

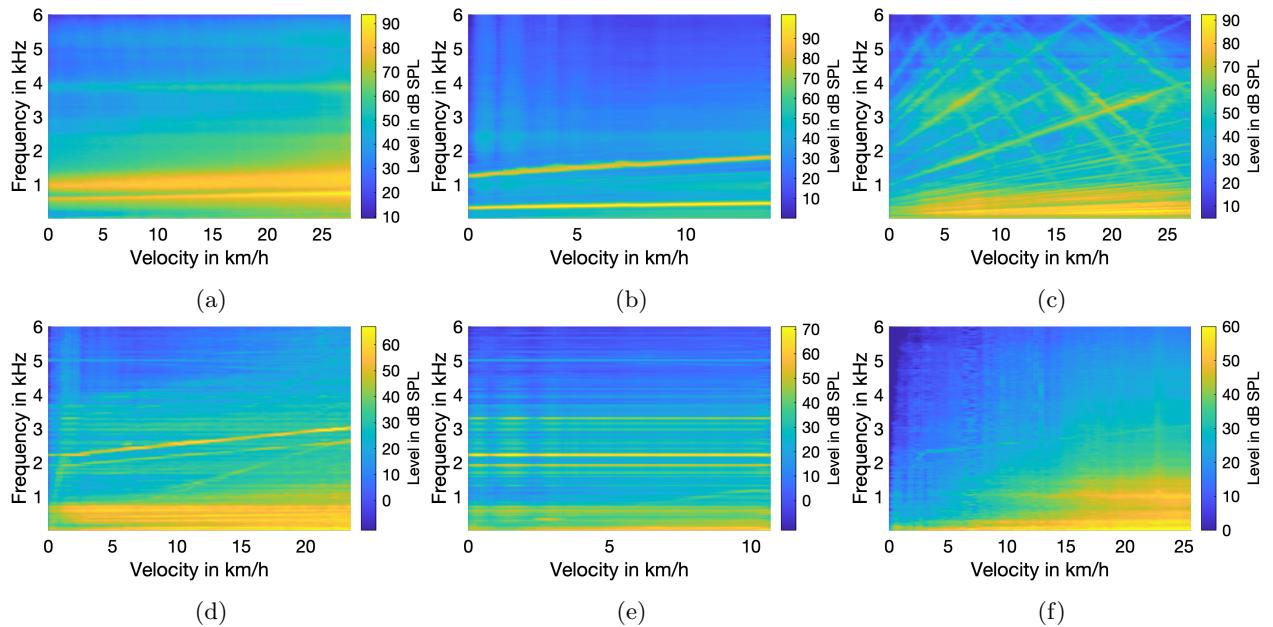


Figure 2: Measured velocity dependent magnitude spectra  $|H(f, v)|$  from microphones mounted on moving vehicle: vehicle A forward AVAS (a), vehicle A backward AVAS (b), vehicle B forward and backward AVAS (c), vehicle C forward AVAS (d), vehicle C backward AVAS (e) and vehicle C tire/road noise (f).

duction of the recorded scenarios and the comparison to their simulated counterparts. The measurements were conducted with a sampling frequency of 48 kHz and on a road with dense asphalt concrete surface under dry and windless conditions. Several passages with different velocity profiles up to 30 km/h were recorded for each vehicle, driving both forward and backward. All measured data is openly accessible at [doi.org/10.5281/zenodo.10610490](https://doi.org/10.5281/zenodo.10610490).

## 2.2 Measurement Results

The following section presents the measured AVAS signals for the three evaluated vehicles as well as the results of the tire/road noise recordings.

### 2.2.1 AVAS Signals

**Figure 2** shows the measurement results of the microphone placed in front of the AVAS loudspeaker for all three vehicles. For these plots, the measured time signals were downsampled to a sampling rate of 12 kHz and divided into overlapping blocks of 512 samples. Each block was assigned a velocity value based on the GPS recordings, transformed into a magnitude spectrum, and the spectra for blocks with similar velocity values were averaged. This results in frequency over velocity plots, which visualize the characteristic velocity dependency of the AVAS signals. Thereby, the three measured electric vehicles differ quite substantially, as described in the following. We recommend listening to the example sounds provided at [doi.org/10.5281/zenodo.10610490](https://doi.org/10.5281/zenodo.10610490) for a better understanding of the different signals.

**Vehicle A** As shown in [Figure 2a](#), vehicle A radiates a band-pass filtered noise in combination with a lower, more narrow-band component when driving forward. The center frequency, as well as the bandwidth of both components, increases with vehicle velocity. When driving backward, vehicle A radiates two highly tonal components, which also increase in frequency with velocity (cf. [Figure 2b](#)). Listening to the recorded signal reveals a clearly audible amplitude modulation, which is also visible in the form of sidebands when analyzing the corresponding frequency over time spectrogram (not shown here for reasons of conciseness).

**Vehicle B** Vehicle B radiates the same AVAS sound in both driving directions, consisting of a large number of different tonal components which partly seem to have a harmonic relation to each other as shown in [Figure 2c](#). Most of those tones increase in frequency when accelerating, others decrease in frequency, and some are strongly amplitude-modulated. Of all measured AVAS signals, we perceive vehicle B as the most “chaotic”; it is also the only signal that contains strong tonal components above 3 kHz. Overall, the signal reminds us more of a science-fiction spaceship sound than the noise expected of a medium-sized passenger car. Whether or not this perceptual discrepancy has consequences for factors such as annoyance or localization accuracy compared to more conservative AVAS sounds is an example for a future study that could be performed using the methods presented in this paper.

**Vehicle C** The forward AVAS of vehicle C consists of both high-frequency tonal components and lower-frequency band-pass filtered noise components (cf. Figure 2d). Compared to, for example, vehicle B, individual tonal components are perceptually not as pronounced and partly masked by background noise. When driving backward, vehicle C emits a recurring “pling” sound with a duration of 1 s per repetition that appears to be independent of the vehicle velocity as shown in Figure 2e.

Summarized, one can divide the evaluated AVAS signals into three categories: signals that mainly consist of a number of tones, signals that mainly consist of band-pass filtered noise components, and signals that consist of a repeating sound. Section 3.1 presents methods to synthesize those three different signal types for arbitrary vehicle velocities based on the presented reference measurements.

### 2.2.2 Tire/Road Noise

Figure 2f shows the measured tire/road noise spectrum for vehicle C. The results for vehicles A and B were found to be very similar and are, for reasons of conciseness, not presented here. It can be seen that the spectrogram contains a pronounced pressure maximum centered around 1 kHz above a velocity of approximately 15 km/h. This pressure maximum is characteristic of tire/road noise [15] and increases in amplitude as well as slightly broadens in frequency range for higher vehicle speeds. Additionally, the spectrogram shows strong low-frequency components that increase in amplitude and upper-frequency limits for higher velocities. These components are assumed to correspond to wind-induced noise in the microphone [17], which are, due to the microphone positioning, expected to be more pronounced for the tire noise measurements than for the AVAS measurements where the car body partly shielded the microphone.

## 3 Auralization

The first step in auralizing a vehicle passage is to accurately synthesize all relevant source signals, which are, for the scope of this paper, limited to the AVAS signal as well as tire/road noise. As the previously described measurements showed, the three evaluated vehicles use substantially different AVAS signal types, requiring the implementation of different synthesis techniques as described in Section 3.1. Besides the source signal, the radiation characteristics of the different sources need to be taken into account, e.g., the sound radiation from an AVAS loudspeaker mounted to the front bumper of a car will most likely be strongest in the forward direction, and tire/road noise typically shows a strong, frequency dependent, amplification perpendicular to the contact surface due to the horn

effect [18]. Section 3.2 presents a boundary element approach to numerically estimate the AVAS radiation directivity for all three evaluated vehicles and uses previously performed tire radiation measurements to construct a generic tire/road noise directivity pattern. All source directivities are encoded into spherical harmonic coefficients to simplify the subsequent processing. Once all source signals and radiation directivities are known, the sound propagation from all sources to a receiver position is modeled using spherical harmonic extrapolation while additionally taking into account vehicle movement, air attenuation, and binaural hearing as described in Section 3.3. The entire auralization model was implemented in Matlab R2023a; all relevant code is available under MIT license at [github.com/leonpaulmueller/evat](https://github.com/leonpaulmueller/evat).

### 3.1 AVAS and Tire/Road Noise Synthesis

Based on the evaluation of the measured reference signals described in Section 2, we differentiate between three different synthesis methods for generating AVAS and tire/road noise signals: (i) subtractive synthesis, where broadband noise is filtered with vehicle velocity-dependent band-pass filters, (ii) additive synthesis, where sine wave oscillators are added up and modulated in amplitude and frequency depending on the vehicle velocity and (iii) sample-based synthesis, where a sample, i.e. a short sound recording, gets repeatedly played back and modified based on the vehicle velocity. The following sections describe methods to analyze the recorded reference signals and resynthesize them using those three approaches. While this paper focuses on recreating measured AVAS signals of existing vehicles, the synthesis methods are designed in a way that allows for creating arbitrary new AVAS sounds.

#### 3.1.1 Subtractive Synthesis

The subtractive synthesis model used in this paper assumes that the output signal, i.e., the AVAS or tire/road noise, can be described by convolving Gaussian white noise with a set of vehicle velocity-dependent impulse responses. The process of determining these impulse responses and subsequently generating a new signal can be divided into an analysis and a synthesis stage described in the following.

**Analysis** In order to determine a set of velocity-dependent filter functions describing the characteristics of a given AVAS or tire/road noise recording, the recorded pressure signal is divided into blocks of  $N$  samples, and each block is assigned a mean velocity value based on the recorded vehicle velocity signal. Those blocks are then transformed into magnitude

309 spectra and averaged depending on the desired ve-  
 310 locity resolution. Magnitude spectra for missing ve-  
 311 locity values are linearly interpolated, and a smooth-  
 312 ing function is applied to avoid discontinuities be-  
 313 tween individual velocity values, resulting in the fi-  
 314 nal matrix of velocity-dependent magnitude spectra  
 315  $|H(f, v)|$  (c.f. Figure 2).

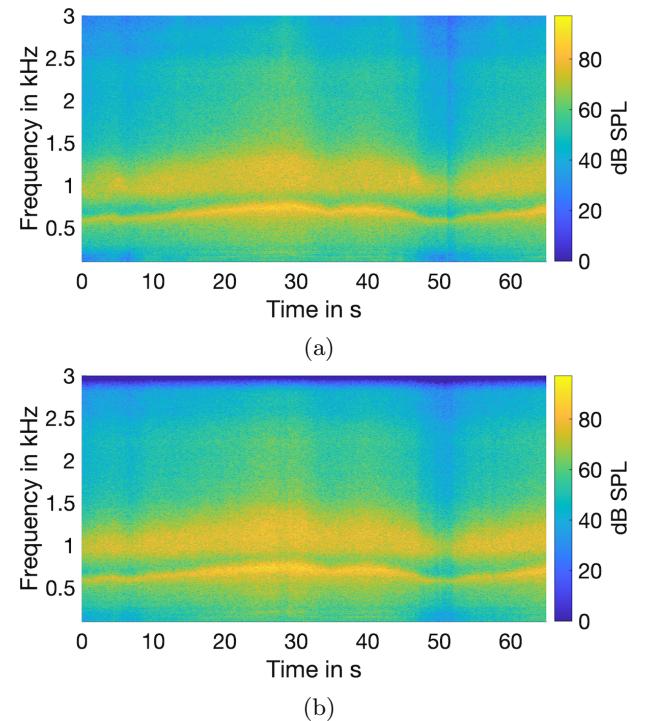
316 **Synthesis** To synthesize a signal for an arbitrary  
 317 vehicle velocity, the new velocity signal is averaged  
 318 over blocks of  $N$  samples. Since the magnitude spec-  
 319 tra  $|H(f, v)|$  do not contain any phase information,  
 320 a minimum phase representation is computed using  
 321 the real cepstrum [19], resulting in a set of velocity-  
 322 dependent impulse responses  $h(n, v)$ . This allows  
 323 block-wise computing of the output signal by con-  
 324 volving the impulse response corresponding to each  
 325 block's velocity value with a block of Gaussian white  
 326 noise. Additionally, the same noise block is convolved  
 327 with the impulse response corresponding to the pre-  
 328 vious velocity value, and both results are cross-faded  
 329 to avoid clicking artifacts between blocks.

330 **Results** The subtractive synthesis approach was  
 331 used to re-synthesize the forward AVAS signals of ve-  
 332 hicles A and C and all tire/road noise signals. Figure  
 333 3 shows an exemplary comparison between the  
 334 measured and the re-synthesized forward AVAS sig-  
 335 na of vehicle A. Thereby, it can be seen that the  
 336 synthesized signal resembles the measured signal very  
 337 well up to 3 kHz, which was set as the upper-frequency  
 338 limit for this synthesis. A perceptual validation of the  
 339 resulting auralization is presented in Section 5, and  
 340 both signals can be listened to via the supplementary  
 341 online repository.

### 342 3.1.2 Additive Synthesis

343 Additive synthesis describes the process of adding  
 344 up multiple sinusoidal signals to create a complex  
 345 sound [20, 21]. Thereby, the individual signal com-  
 346 ponents are typically generated by individual oscilla-  
 347 tors, which, independent of each other, may have a  
 348 time-varying amplitude or frequency. This can either  
 349 result in relatively simple sounds, such as the back-  
 350 ward AVAS of vehicle A consisting of two tones that  
 351 increase their pitch with velocity (c.f. Figure 2b), or  
 352 in more complex sounds such as the vehicle B AVAS,  
 353 consisting of a large number of different tones that are  
 354 partly in harmonic relation to each other (c.f. Fig-  
 355 ure 2c). Figure 4 gives a high-level overview of the  
 356 implemented additive analysis and synthesis model  
 357 described in the following.

358 **Synthesis** We assume that the desired signal con-  
 359 sists of the sum of  $U$  different simple harmonic oscil-  
 360 lators where the  $u$ -th oscillator at time sample  $n$  has  
 361 an amplitude of  $A_u(n)$  and a frequency  $f_u(n)$  that



362 Figure 3: Vehicle A forward AVAS measurement (a)  
 363 and synthesized forward AVAS signal (b).  
 364

365 both change depending on the vehicle velocity  $v(n)$ .  
 366 Furthermore, the amplitude of each oscillator is mod-  
 367 ulated by an additional time-varying harmonic oscilla-  
 368 tion with amplitude  $\check{A}_u(n)$  and frequency  $\check{f}_u(n)$ . This  
 369 means that an individual oscillator is characterized by  
 370 the four parameters  $A_u(n)$ ,  $f_u(n)$ ,  $\check{A}_u(n)$  and  $\check{f}_u(n)$ ,  
 371 which all depend on the vehicle velocity. For each  
 372 oscillator, the relation between those parameters and  
 373 the velocity is described by four sets of polynomial co-  
 374 efficients  $C_{u,q}^f$ ,  $C_{u,q}^A$ ,  $C_{u,q}^{\check{f}}$  and  $C_{u,q}^{\check{A}}$  which can either be  
 375 manually set to design an arbitrary new AVAS signal  
 376 or can be obtained from analyzing an existing AVAS  
 377 recording as described in the next paragraph. This  
 378 means that, for example, the frequency of the  $u$ -th  
 379 oscillator at time sample  $n$  can be calculated from  
 380 the polynomial coefficients  $C_{u,q}^f$  with degree  $Q$  and  
 381 the vehicle velocity  $v(n)$  as  
 382

$$383 f_u(n) = \sum_{q=0}^Q C_{u,q}^f v(n)^q. \quad (1)$$

384 Since the modulation signal for each oscillator is  
 385 expected to change in amplitude and frequency over  
 386 time and the phase argument of such a time-varying  
 387 signal is proportional to the integral of the instantane-  
 388 ous frequency [22, Sec. 5.6], the modulation signal  
 389  $\check{s}_u(n)$  can be constructed as

$$390 \check{s}_u(n) = \check{A}_u(n) \cos \left( \frac{2\pi}{f_s} \sum_{i=0}^n \check{f}_u(i) \right). \quad (2)$$

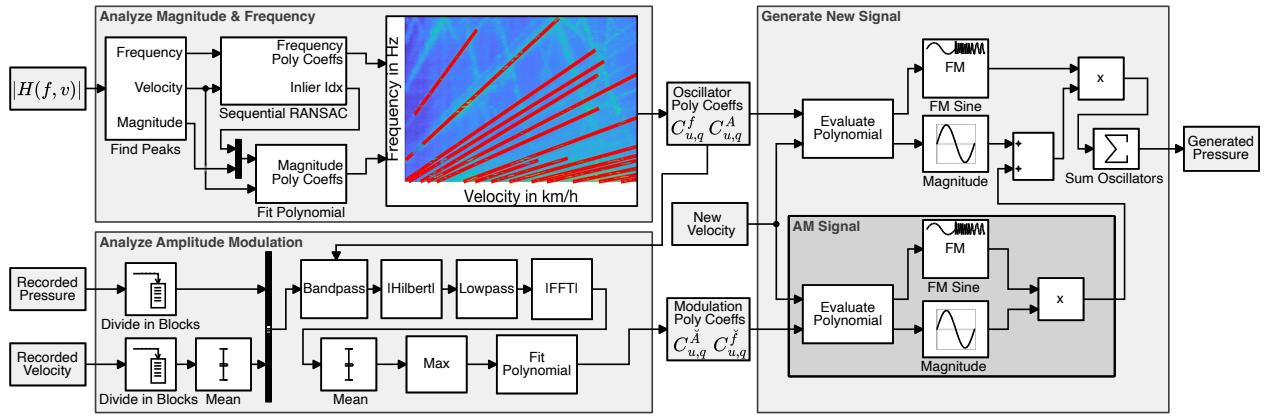


Figure 4: Additive analysis and synthesis model. The red lines in the spectrogram visualize RANSAC results for vehicle B, i.e., amplitude and frequency of individual oscillators. The velocity-dependent magnitude spectrum  $|H(f, v)|$  is obtained from the analysis stage of the subtractive synthesis method described in Section 3.1.1.

385 The output signal  $s(n)$  then corresponds to the sum  
 386 of all  $U$  oscillators, which each consist of a frequency-  
 387 modulated sinusoidal multiplied with the corresponding  
 388 amplitude modulation signal:

$$389 s(n) = \sum_{u=1}^U (A_u(n) + \check{s}_u(n)) \cos \left( \frac{2\pi}{f_s} \sum_{i=0}^n f_u(i) \right). \quad (3)$$

390 **Analysis** To determine the coefficient sets  $C_{u,q}^f$ ,  
 391  $C_{u,q}^A$ ,  $C_{u,q}^{\tilde{f}}$  and  $C_{u,q}^{\tilde{A}}$  from a recorded AVAS signal, the  
 392 same velocity dependent magnitude spectra  $|H(f, v)|$   
 393 used in Section 3.1.1 is analyzed to find peaks with  
 394 a user-defined prominence, threshold and inter-peak  
 395 distance for each velocity value. Those peaks' am-  
 396 plitude and frequency values are then processed us-  
 397 ing a sequential random sample consensus (RANSAC)  
 398 approach [23, 24], which iteratively analyzes the ob-  
 399 served peak values to find  $U$  different subsets best  
 400 described by least-squares fitting a polynomial with  
 401 degree  $Q$ . This results in two sets of polynomials,  
 402  $C_{u,q}^f$  and  $C_{u,q}^A$ , describing the velocity-dependent fre-  
 403 quency and amplitude behavior of all oscillators. The  
 404 red lines in the spectrogram shown in Figure 4 vis-  
 405 ualize those polynomials, where each line represents an  
 406 individual oscillator. As can be seen, the estimated  
 407 polynomials cover most but not all of the tonal com-  
 408 ponents included in the analyzed AVAS signal.

409 Additionally, the strength and frequency of the am-  
 410 plitude modulation for each oscillator are analyzed by  
 411 block-wise band-pass filtering the recorded AVAS sig-  
 412 nal according to the previously determined frequency  
 413 for each oscillator. Taking an FFT of the Hilbert enve-  
 414 lope of those band-pass filtered signals then allows de-  
 415 termining the velocity-dependent amplitude modula-  
 416 tion frequency and strength which are then described  
 417 by the two polynomial sets  $C_{u,q}^{\tilde{f}}$  and  $C_{u,q}^{\tilde{A}}$ .

**Results** Figure 5 shows a time-frequency represen-  
 418 tation of the measured and additively re-synthesized  
 419 AVAS signal for vehicle B. It can be seen that the gen-  
 420 erated signal successfully reproduced a large number  
 421 of tonal components contained in the original record-  
 422 ing, but not all of them. Additionally, both signals dif-  
 423 fer in background noise level as the synthesized signal  
 424 solely consists of pure tones while the recorded sig-  
 425 nal contains a broadband background noise of around  
 426 30 dB. However, in this specific case, the background  
 427 noise most likely corresponds to wind-induced noise  
 428 in the microphone and cross-talk from tire/road noise;  
 429 hence, it is assumed not to be part of the AVAS sig-  
 430 nal and should not be included in the source signal  
 431 synthesis. However, other AVAS signals might con-  
 432 tain strong tonal and broadband noise components,  
 433 requiring a combination of additive and subtractive  
 434 synthesis.

### 3.1.3 Sample-Based Synthesis

In contrast to subtractive and additive synthesis,  
 436 sample-based synthesis uses a pre-recorded sound,  
 437 a so-called sample, instead of simple sinusoids or  
 438 noise as a synthesis foundation. Different variations  
 439 of sample-based synthesis have been previously used  
 440 to, for example, auralize combustion engine noise [25,  
 441 26], using sophisticated algorithms to modify sample  
 442 properties such as pitch or time scale. However, for  
 443 this paper, only the backward AVAS of vehicle C,  
 444 consisting of a simple “plinging” sound played back  
 445 repeatedly with a constant pitch, is of interest for  
 446 a sample-based synthesis approach. Therefore, the  
 447 implemented method was limited to only modulating  
 448 sound pressure level and repetition rate depending on  
 449 the vehicle velocity.

**Synthesis** The implemented sample-based synthe-  
 451 sis model constructs an output signal based on rep-  
 452

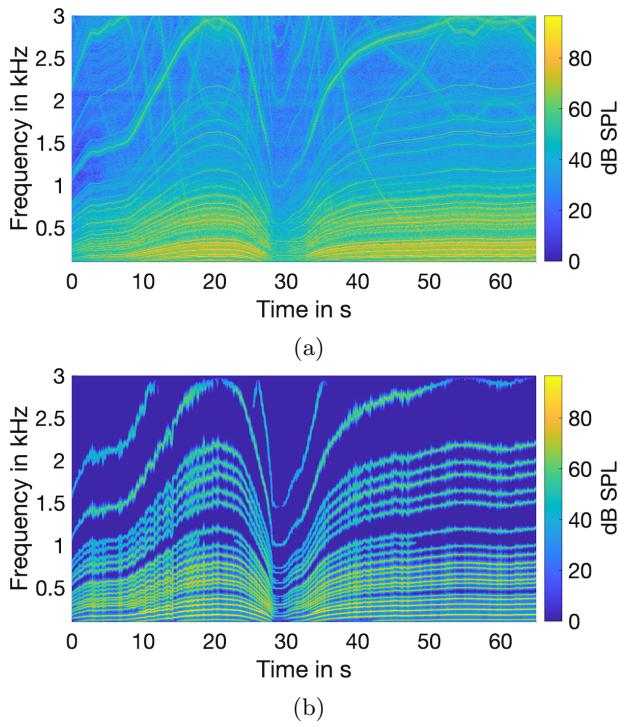


Figure 5: Vehicle B AVAS measurement (a) and additively synthesized AVAS signal (b).

etitions of a pre-recorded sound sample. Thereby, both the RMS sound pressure level and the repetition rate are assumed to be velocity dependent and are described by a set of polynomial coefficients, similar to the polynomials used for the additive synthesis approach in Section 3.1.2. To synthesize an AVAS signal for an arbitrary new velocity, the instantaneous repetition rate and sound pressure level are calculated from those polynomials according to Equation (1) and then used to distribute the pre-recorded samples in an output signal vector as well as scale them to achieve the desired RMS-values.

**Analysis** Analyzing a recorded AVAS signal for sample-based synthesis requires manually selecting one period of the desired sound sample in the recording. The entire recording is then analyzed using autocorrelation to find repetitions of the selected sound. The RMS value of each repetition and the spacing between repetitions is assigned to the corresponding recorded vehicle velocity value, which allows for the fitting of polynomials describing the velocity dependency of both parameters. Signal repetitions correlating strongly with the selected reference sound are then averaged in the time domain to obtain a clean signal sample and suppress potential background noise.

**Results** Figure 6 shows a comparison between the recorded and the re-synthesized sample-based backward AVAS signal of Vehicle C. Both signals appear to be very similar both in time and frequency struc-

ture, apart from the fact that the synthesized signal shows a lower overall background noise level. However, similarly to the additive synthesis results, we assume that the background noise in the recording is rather an artifact than part of the AVAS signal, which means that the lower noise level in the generated signal is beneficial for auralization purposes.

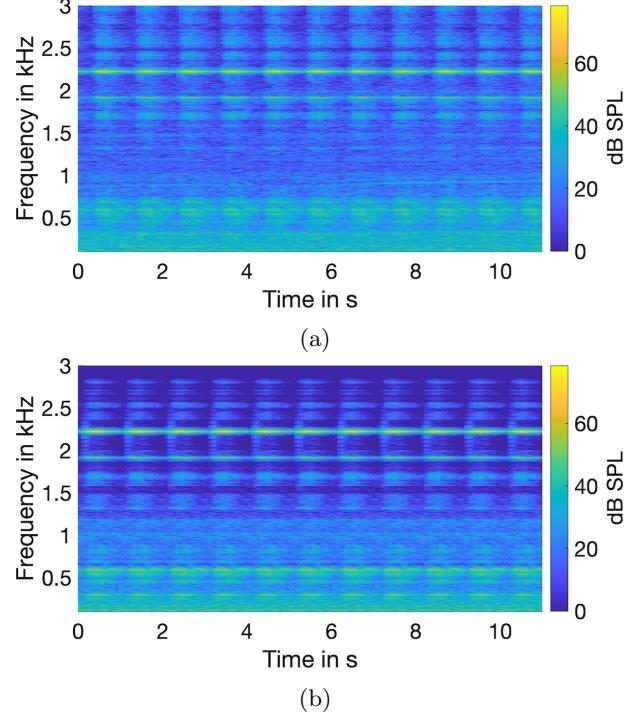


Figure 6: Vehicle C backward AVAS measurement (a) and synthesized backward AVAS signal (b).

### 3.1.4 Source Signal Synthesis Conclusion

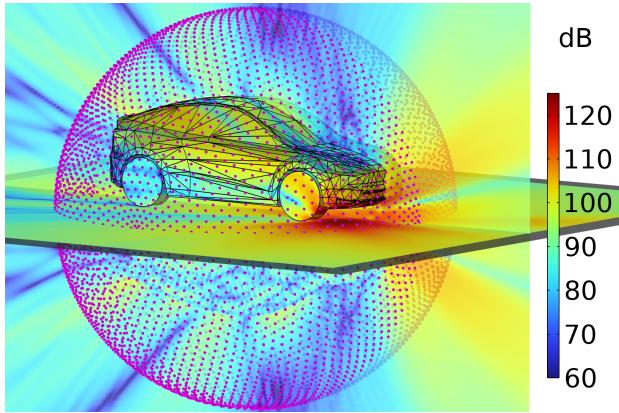
Based on the characteristics of the measured AVAS and tire/road noise signals, three different source signal synthesis methods have been implemented. All three methods have been shown to work well for the signals generated by the evaluated vehicles. However, additional fine-tuning might be required when adapting the methods to synthesize other types of signals. Thereby, the choice of the method should, of course, be determined by the character of the signal of interest.

## 3.2 Directivity

For the auralization approach presented in this paper, we assume that the sound radiated by an electric vehicle is a superposition of two types of sources: tire/road noise and the AVAS warning sound. While the previous section described the synthesis of the corresponding source signals, one also has to consider sound radiation properties, which are expected to be both frequency and space-dependent. Whether or not

509 this spatial directivity is perceptually relevant could  
 510 be a subject of a future study; to allow such research,  
 511 this work aimed to reproduce the source directivities  
 512 as accurately as possible. Therefore, the AVAS radi-  
 513 ation patterns for all three vehicles were numerically  
 514 estimated using the boundary element method (BEM)  
 515 as described in [Section 3.2.1](#), and the tire/road noise  
 516 directivity was modeled based on previous measure-  
 517 ments (c.f. [Section 3.2.2](#)). Both AVAS and tire di-  
 518 rectivities were encoded using spherical harmonic ex-  
 519 pansion as described in [Section 3.2.3](#) to allow more  
 520 straightforward spatial processing and propagation  
 521 modeling.

### 522 3.2.1 AVAS Directivity



523 Figure 7: Simplified 3D model of vehicle A with BEM  
 524 results for radiated sound pressure at  $f = 2\text{ kHz}$  and  
 525 evaluation points on Lebedev grid. The mirrored  
 526 pressure below the ground plane is a consequence of  
 527 the symmetry boundary condition used to model an  
 528 infinite sound-hard ground.

529 **Methods** In order to obtain a numerical estimate  
 530 of the AVAS radiation directivity, all three vehicles  
 531 were modeled using the boundary element method  
 532 (BEM) in COMSOL MULTIPHYSICS 6.1. The vehi-  
 533 cle geometries were based on simplified, commercially  
 534 available 3D models of the individual cars where the  
 535 AVAS loudspeaker was substituted by a single disk  
 536 with a 5 cm radius embedded in the vehicle chassis.  
 537 This disk was then excited with a velocity of  $\frac{1}{j\omega}\text{ m/s}$ ,  
 538 which means that the radiated sound pressure would  
 539 be constant over all frequencies if the source would  
 540 be a monopole in free field. An infinite, sound-hard  
 541 ground was included in the simulation by introducing  
 542 a symmetry boundary condition. A simple, porous  
 543 absorber impedance model was assigned to the vehi-  
 544 cle floor to avoid numerical problems caused by  
 545 resonances between the vehicle and the ground. The  
 546 resulting sound pressure was then evaluated at  
 547 5810 points of a 131st order Lebedev sphere [27] with 3 m  
 548

549 radius surrounding the vehicle as visualized in [Fig-  
 550 ure 7](#). Since the complex sound pressure on a surface  
 551 enclosing all sources is known, this pressure can be ex-  
 552 trapolated to any position outside of the sphere [28],  
 553 in this case by using spherical harmonic expansion as  
 554 described in [Section 3.2.3](#). Thereby, the fact that the  
 555 introduced symmetry boundary condition leads to a  
 556 mirrored pressure field below the ground plane au-  
 557 tomatically results in correct ground reflections when  
 558 extrapolating the pressure from the evaluation sphere.  
 559 This correct extrapolation would not be the case when  
 560 setting the pressure below ground to zero or only  
 561 evaluating the upper half-sphere; one can also inter-  
 562 pret this approach as introducing an additional image  
 563 source below the ground.

564 The model was solved up to 3 kHz in 30 Hz steps,  
 565 which, when transformed to time domain, results in  
 566 impulse responses describing the propagation from  
 567 the AVAS loudspeaker to receiver positions on the  
 568 evaluation grid with a sampling rate of 6 kHz and a  
 569 duration of 33.3 ms. This upper-frequency limit was  
 570 set due to high computational demands and effectively  
 571 limits the maximum possible frequency for the AVAS  
 572 signal auralization to 3 kHz. Since none of the mea-  
 573 sured AVAS signals except for vehicle B showed sig-  
 574 nificant contents above 3 kHz and higher frequency  
 575 BEM calculations would have been infeasible with the  
 576 available computational resources, it was decided to  
 577 accept this limitation for the purposes of this study.  
 578 When higher frequency radiation patterns are needed,  
 579 the BEM model could be solved using more computa-  
 580 tional power or be extended by a less computationally  
 581 demanding approach for higher frequencies, such as  
 582 ray tracing.

583 **Results** [Figure 8](#) shows polar representations of the  
 584 calculated AVAS radiation directivity for vehicle A.  
 585 These results show that the radiation in the horizontal  
 586 plane is focused towards  $330^\circ$  and that the sound ra-  
 587 diation in the frontal plane is also skewed towards this  
 588 direction for all evaluated frequency bands. This di-  
 589 rectivity appears reasonable as the AVAS loudspeaker  
 590 of this vehicle is mounted on the right side of the front  
 591 bumper (c.f. [Figure 1d](#)). While the details of the ra-  
 592 diation patterns obtained from the BEM calculations  
 593 might not be perfectly accurate due to, e.g., devia-  
 594 tions in the vehicle geometry and unknown material  
 595 properties, they strongly indicate that this specific ve-  
 596 hicle radiates the AVAS signal mostly to the front  
 597 right relative to its driving direction and less to the  
 598 back and to the left side. This observation was also  
 599 confirmed by in-situ directivity measurements of the  
 600 vehicle A AVAS system, which we, in order to limit  
 601 the extent of this paper, do not describe in further de-  
 602 tail. The radiation directivities calculated for vehicles  
 603 B and C were calculated using the same methods and  
 604 are attached in [Appendix A](#) ([Figure 16](#)). How exactly  
 605 these patterns differ and whether or not they are ben-

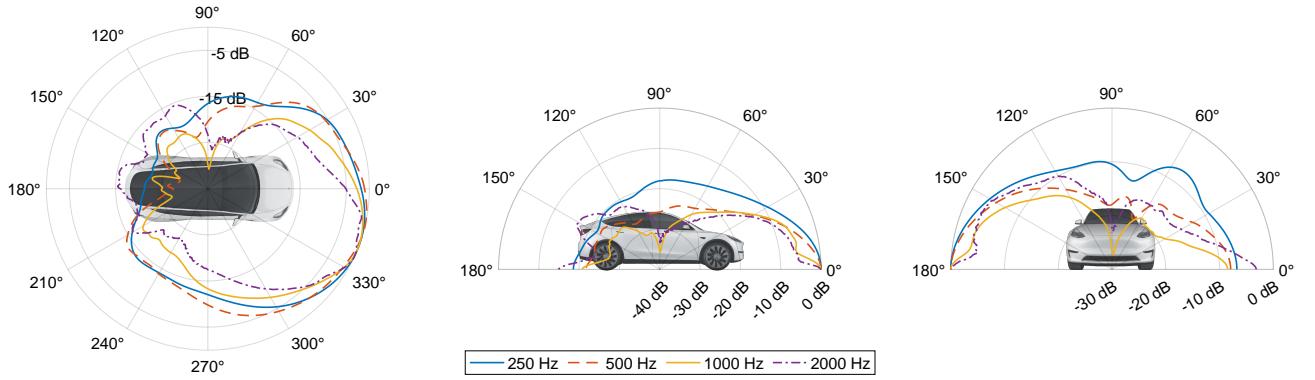
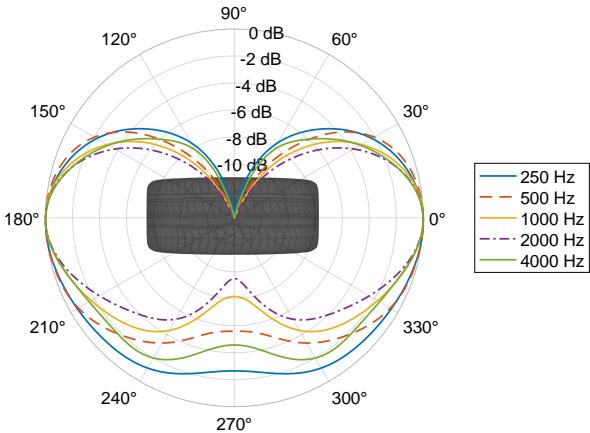


Figure 8: Vehicle A AVAS radiation directivity results of BEM calculation for the horizontal, median, and frontal plane. Normalized to the maximum for each frequency band.

599 official for the intended AVAS warning purposes would  
 600 be a relevant topic for a follow-up study. For the scope  
 601 of this paper, we conclude that the AVAS radiation  
 602 in the relevant frequency range is not omnidirectional  
 603 and should, as long as it is not shown to be perceptually  
 604 irrelevant, be included in the auralization.

### 605 3.2.2 Tire Directivity



637 Figure 9: Tire Directivity in octave bands, normalized  
 638 to the maximum for each band and attenuated  
 639 towards the direction of the vehicle body.

640 The implemented tire directivity is based on mea-  
 641 surements performed in [18]. Thereby, a microphone  
 642 was flush mounted to the ground below the rolling  
 643 surface of a commercial tire (type 155SR13) with a  
 644 distance of 10 mm to the tire/road contact point. The  
 645 tire itself was installed on a middle-class car, and a  
 646 loudspeaker was placed with 7.5 m distance to the cen-  
 647 ter of the tire and 1.2 m above the ground at a hor-  
 648 izontal angle of 0°, 15°, 30°, 45°, 60°, 75° and 90°  
 649 relative to the tire normal axis. Transfer functions  
 650 between the loudspeaker and the microphone were  
 651 measured outdoors on a dense asphalt surface using  
 652 the maximum length sequence technique. Assuming  
 653 reciprocity, these transfer functions describe the ra-

654 diation from the tire/road contact point to the envi-  
 655 ronment, which were then normalized by the transfer  
 656 function measured for 0° and mirrored to construct a  
 657 full sphere 360° radiation pattern. Half of this radi-  
 658 ation directivity was manually attenuated to com-  
 659 pensate for the car body blocking parts of the radiated  
 660 sound, which resulted in a polar pattern as shown in  
 661 Figure 9. For the tires on the opposite side of the car,  
 662 the pattern was rotated by 180°. While the resulting  
 663 directivity is certainly not as accurate as, for example,  
 664 a detailed radiation simulation of the exact tires used  
 665 for the reference pass-by measurements, the main fo-  
 666 cus of this work is the accurate auralization of the  
 667 specific AVAS systems more than the exact reproduc-  
 668 tion of the tire/road noise. We, therefore, assume that  
 669 the obtained generic tire/road noise radiation pattern  
 670 is sufficient for this purpose.

### 671 3.2.3 Spherical Harmonic Expansion

672 The previously described radiation directivities, i.e.,  
 673 the transfer functions from the sound source to a  
 674 spherical grid of evaluation points, can be used to di-  
 675 rectly calculate the sound propagation from a moving  
 676 electric vehicle to a receiver position. However, doing  
 677 so is not very convenient, especially when consider-  
 678 ing that other researchers might want to embed the pre-  
 679 sented radiation directivities in their own auralization  
 680 tools. One complication with directly using the ob-  
 681 tained transfer functions for the auralization is that  
 682 the AVAS and tire directivities have different spatial  
 683 resolutions, and the exact coordinates of the evalua-  
 684 tion points need to be known for further processing.  
 685 Additionally, calculating the directivity for a polar  
 686 angle that does not lie on the evaluation grid requires  
 687 some form of interpolation, and for real-time auraliza-  
 688 tions, the spatial resolution of the AVAS directivities  
 689 might need to be decreased.

690 All these problems are simplified by encoding the  
 691 radiation patterns using spherical harmonic (SH) ex-  
 692 pansion [29–32]. While already established for vir-  
 693 tual acoustics applications, this concept of handling

complex radiation patterns has also gained entry to other fields such as railway acoustics [33] or road traffic noise auralizations [34]. Thereby, the fundamental approach is that a sound field defined on the surface of a sphere with radius  $r_0$  is decomposed into a sum of orthogonal spherical harmonic basis functions, resulting in a set of spherical harmonic coefficients. This set of SH coefficients then allows the extrapolation of the pressure at an arbitrary angle and distance as well as the reduction of the spatial resolution by truncating the order of the spherical harmonics. The following paragraphs describe the implemented spherical harmonic encoding and extrapolation methods and evaluate how well the AVAS radiation patterns are reproduced using these techniques.

**Encoding** If  $p(r_0, \phi, \theta, \omega)$  represents the pressure on an observation sphere with radius  $r_0$ , which, in our case, corresponds to the BEM calculations and tire radiation measurement results, this pressure can be expanded into a set of spherical harmonic expansion coefficients  $W_l^m(r_0, \omega)$  as [28, Ch. 6.3.3]

$$p(r_0, \phi, \theta, \omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^l W_l^m(r_0, \omega) Y_l^m(\phi, \theta). \quad (4)$$

Thereby,  $Y_l^m(\phi, \theta)$  represents the SH basis functions with order  $l$  and degree  $m$  for the azimuth angle  $\phi$  and the colatitude angle  $\theta$ . When limiting the maximum SH order to  $l = L$  and considering spatially discrete pressure observation points, Equation (4) can be written in vector-matrix form as [32]

$$\mathbf{p} = \mathbf{Y} \mathbf{W}. \quad (5)$$

If the number of discrete observation points in  $\mathbf{p}$  is greater than or equal to  $(L+1)^2$ , Equation (5) can be solved in a least-squares sense to obtain the SH coefficient matrix  $\mathbf{W}$  [35]. The maximum SH order  $L$  determines the spatial resolution of the encoded directivities; an advantage of the spherical harmonic encoding is that the spatial resolution can be reduced without any additional interpolation or down-sampling by simply truncating  $W_l^m(r_0, \omega)$  to a lower SH order. Based on the number of pressure observation points obtained from the BEM calculations and the tire measurements, the AVAS directivities were encoded with SH order  $L = 64$  and the tire directivity with  $L = 16$  which, based on studies such as [36, 37] and the perceptual validation performed in Section 5, is expected to be sufficient for auralization purposes.

**Extrapolation** Assuming that the reference pressure was observed on a sphere with radius  $r_0$ , the pressure  $p(r, \phi', \theta', \omega)$  at any position with  $r \geq r_0$  can be extrapolated by multiplying the SH coefficients,  $W_l^m(r_0, \omega)$ , and the SH basis functions for those new

positions,  $Y_l^m(\phi', \theta')$ , and scaling the result with the  $l$ -th order spherical Hankel function of the first kind,  $h_l(kr)$ , as [28, Eq. 6.94]

$$p(r, \phi', \theta', \omega) = \sum_{l=0}^L \frac{h_l(kr)}{h_l(kr_0)} \sum_{m=-l}^l W_l^m(r_0, \omega) Y_l^m(\phi', \theta'). \quad (6)$$

For the AVAS directivities, the symmetry condition embedded in the BEM model results in a mirrored pressure field below the ground (c.f. Section 3.2.1), which, when included in the SH expansion, leads to the extrapolation method correctly reproducing all ground reflections. For the tire/road noise, we assume that the sound source corresponds to the tire contact point on the ground; hence, there are no first-order ground reflections. Second-order ground reflections, such as tire/road noise scattered from the tire or vehicle chassis to the ground, are not correctly extrapolated using the measured tire directivities.

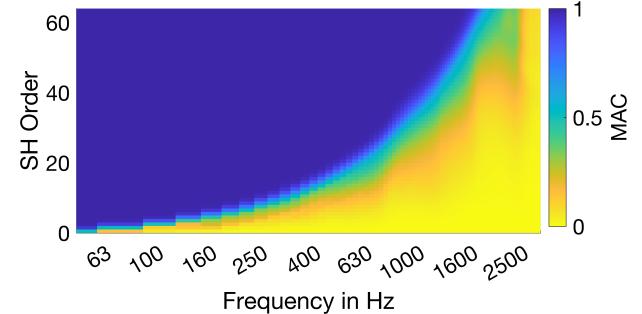


Figure 10: Modal assurance criterion between SH directivities extrapolated to validation grid and corresponding BEM results as a function of SH order and frequency.

**Validation** To evaluate the accuracy of the SH encoding and extrapolation method, the modal assurance criterion (MAC) [38] between the pressure extrapolated from the SH coefficient set with  $r_0 = 3$  m to a 2861-point spherical validation grid with  $r = 6$  m and the BEM results for the same validation positions was calculated as

$$\text{MAC}(\omega) = \frac{|\mathbf{p}_{\text{SH}}^H(\omega) \mathbf{p}_{\text{BEM}}(\omega)|^2}{(\mathbf{p}_{\text{BEM}}^H(\omega) \mathbf{p}_{\text{BEM}}(\omega)) (\mathbf{p}_{\text{SH}}^H(\omega) \mathbf{p}_{\text{SH}}(\omega))}. \quad (7)$$

Thereby,  $\mathbf{p}_{\text{SH}}$  and  $\mathbf{p}_{\text{BEM}}$  correspond to the complex pressure values obtained from the SH extrapolation and the BEM simulation; the superscript H marks the Hermitian transpose. Similar to a correlation coefficient, the MAC describes the degree of linearity between the extrapolated SH pressure and the pressure at the validation grid obtained directly from BEM. A MAC value of 1 indicates a perfectly linear spatial

dependency between both pressure sets, which, compared to spatially independent measures such as the RMS error, gives more meaningful insights into the actual similarity of the radiation patterns. This comparison allows estimating how well the SH extrapolation method can reproduce pressure at positions other than the ones used to calculate the SH directivities.

Figure 10 shows the MAC results as a function of SH order and frequency. It can be seen that higher frequencies, in general, require higher SH orders to reproduce the AVAS radiation pattern correctly. At the maximum evaluated SH order of  $L = 64$ , both pressure sets have a perfectly linear dependency up to a frequency of 1600 Hz. This result indicates that the SH extrapolation correctly reproduces the direct pressure radiated by the AVAS speaker as well as captures the ground reflections and scattering on the vehicle body included in the BEM simulations. Around 2500 Hz, the MAC for  $L = 64$  drops to zero, indicating a significant error between the SH extrapolated data and the BEM results. This comparably large error for low SH orders could be a consequence of the acoustic center of the sound source not being placed in the center of the evaluation sphere [39], which results in a more complicated pressure pattern on the evaluation sphere and hence requires a higher SH order than when the evaluation sphere would be aligned with the acoustic center of the sound source. However, since scattering from the vehicle has to be considered part of the sound source, the entire car must be enclosed by the evaluation sphere. Centering the sphere on the AVAS speaker mounted in the front bumper would hence require a larger radius  $r_0$ , resulting in a larger minimum auralization distance, making the method unsuitable for the auralization of close-distance vehicle passages.

While this upper-frequency limit is sufficiently high to cover the most relevant parts of the measured AVAS signals, the MAC validation indicates that not all AVAS components are reproduced correctly without increasing the SH order even further. However, increasing the SH order above  $L = 64$  would require an even denser evaluation grid and is additionally complicated by the limited numerical resolution when calculating the SH basis functions in Matlab. Fortunately, several perceptual studies have shown that high-frequency deviations in sound source directivity are often not audible and that much lower SH orders may be sufficient for auralization purposes [36, 37]. Additionally, the modal assurance criterion evaluates the similarity of both magnitude and phase, whereas, in practice, it is often considered acceptable only to reproduce the correct magnitude of sound source directivities, significantly reducing the required SH order [32]. Since the overall goal of this work was to create perceptually accurate rather than numerically perfect simulations, it was concluded that an AVAS SH order of  $L \leq 64$  is sufficient for this purpose. This

assumption was further investigated in a listening experiment presented in Section 5.

### 3.3 Propagation

The movement of the outdoor sound source was implemented using the concept of moving Green's functions [40]. This means that the desired trajectory of the vehicle is spatially discretized according to the desired sampling frequency, i.e., each audio sample is assigned to a corresponding source position. Transfer functions describing the propagation from each of those discrete source positions to the receiver position are then calculated by extrapolating the SH directivities using Equation (6); applying an inverse Fourier transform results in a set of Green's functions  $g_j(n)$  for all discrete source positions  $j$ . The number of source positions  $N$  is equal to the number of output samples and source signal samples; the length of each Green's function depends on the frequency resolution of the SH coefficients. Combining these Green's functions with a source signal  $s(n)$  obtained from the synthesis methods described in Section 3.1 allows for calculating the resulting pressure at the receiver position  $p_{\text{rec}}(n)$  by convolving each Green's function with the corresponding sample of the source signal as

$$p_{\text{rec}}(n) = \sum_{i=0}^{N-1} s(n-i) \cdot g_{n-i}(i). \quad (8)$$

To account for atmospheric absorption, the resulting pressure is filtered according to ISO 9613-1:1993 [41] by attenuating individual third-octave bands depending on the instantaneous distance between the source and the receiver. For a headphone-based reproduction, the pressure signal is then block-wise convolved with head-related transfer functions (HRTFs) corresponding to the instantaneous angle between the source and the receiver, resulting in a binaural output signal. Alternatively, the HRTF for each sample  $n$  can be convolved with the corresponding Green's function to directly obtain a binaural signal from Equation (8). Since the auralization model assumes that the overall vehicle radiation is composed of five different sound sources, i.e., four tires and the AVAS loudspeaker, which all have a different spatial orientation relative to the receiver position, the previously described process is performed separately for all five sound sources and the resulting binaural pressure signals are added up to obtain a summation of all sound sources. Finally, binaural ambient noise recorded at the exact location used for the measurements in Section 2 is added to the output signal to create a more lifelike scenario instead of simulating a vehicle passing by in a perfectly silent environment.

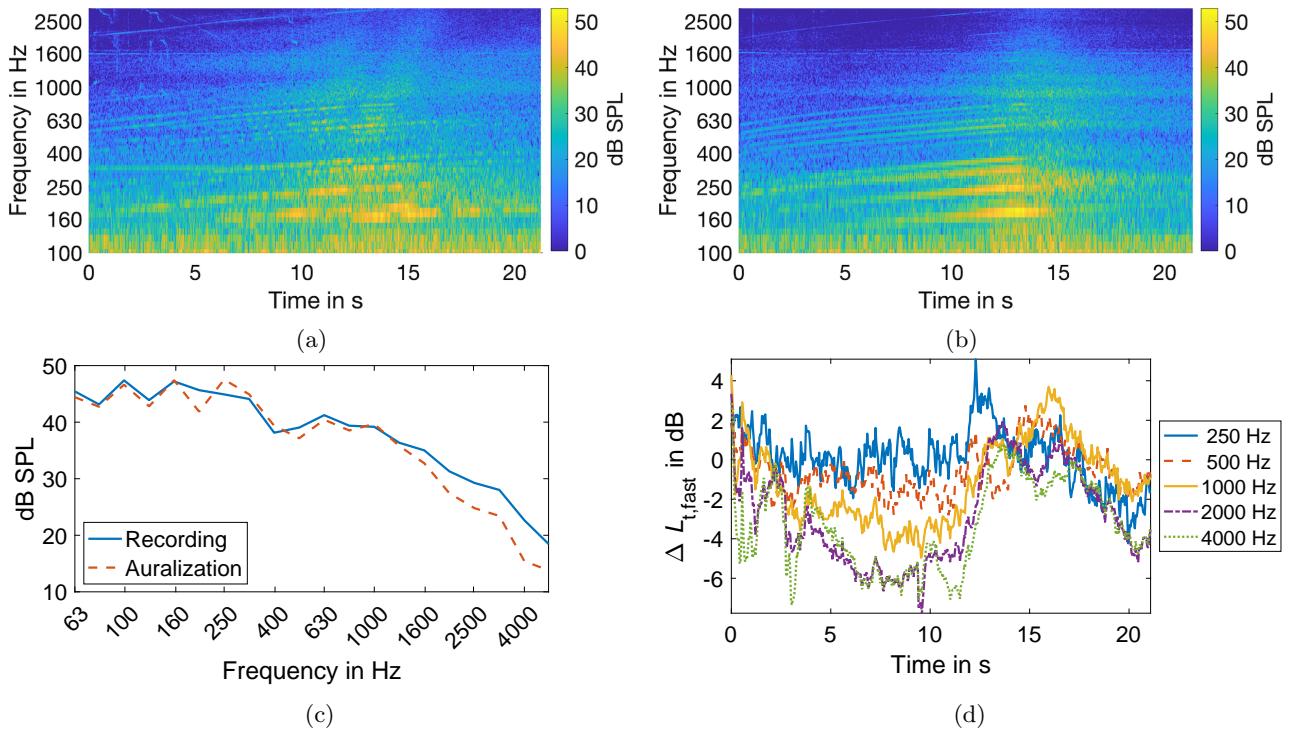


Figure 11: Recorded (a) and auralized (b) passage of vehicle B, third-octave band levels of vehicle B recording and auralization (c) and difference in octave-band fast-weighted levels between vehicle B recording and auralization (d).

## 4 Results

In order to numerically evaluate the quality of the auralization model, several passages of all vehicles were simulated using the velocity profiles of the corresponding recordings. Using the same velocity profile and vehicle type means that, ideally, the auralization should result in a signal identical to the recorded signal. Figure 11 shows an exemplary comparison between a recorded and auralized passage of vehicle B with the same vehicle velocity profile for a roadside observer position. Comparing the spectrograms of the recorded and generated signals (c.f. Figure 11a, Figure 11b) shows that the generated signal has less tonal components than the recorded signal, which is caused by the generated AVAS signal (compare Figure 5). Additionally, the tonal components in the generated signal decrease drastically as soon as the vehicle passes the observer at ca. 15 seconds, while, in the recorded signal, tonal components fade out more slowly. This difference is also evident in the octave band comparison over time shown in Figure 11d. A possible explanation for these deviations might be that the numerically estimated AVAS directivity is inaccurate, i.e., the AVAS radiation is more omnidirectional than the BEM results indicate. This might be caused by the simplified vehicle geometry and the lack of surface roughness and diffuse reflections in the BEM model. Another factor that could lead to such an overly pronounced magnitude change during the

vehicle passage is that the simulated signals, except the ground reflections included in the directivities, assume free field sound propagation. The recordings, however, were made in proximity to buildings, resulting in additional reflections and, hence, a more diffuse sound field at the receiver position. Even if the AVAS radiation is highly directional, additional diffuse reflections would automatically decrease the influence of the source directivity during the passage. Implementing an image source model by mirroring the calculated directivities would allow the inclusion of those reflections in the simulation, potentially resulting in more accurate auralizations at the cost of higher computational demand. Finally, omitting the time structure and comparing both signals in third-octave bands, as shown in Figure 11c, reveals that the auralization relatively accurately reproduces the overall time-averaged sound pressure levels, which might be of interest for research in the context of traffic noise regulations. To summarize, the numerical comparison between the recorded and synthesized passage of vehicle B revealed differences in the time/frequency structure of the signals, which could originate from an incorrect directivity model or a lack of environmental reflections. Similar differences are also visible when comparing auralizations and reference recordings for the other evaluated vehicle as shown in Appendix A (Figure 17).

## 902 5 Perceptual Validation

903 While the previously presented numerical vali-  
 904 dations already revealed that the auralization results  
 905 are no perfect reproductions of the reference record-  
 906 ings, a perceptual validation is necessary to deter-  
 907 mine whether or not those differences are relevant to  
 908 the method's intended purpose of performing AVAS-  
 909 related listening experiments. Thereby, one has to  
 910 first specify the needs of this application to be able  
 911 to decide on which "quality level" is required. For  
 912 example, does the auralization necessarily have to be  
 913 numerically identical to in-situ measurements? Or is  
 914 it sufficient if the auralization is "authentic", i.e., per-  
 915 ceptually indistinguishable in direct comparison to a  
 916 real sound acting as external reference [42, 43]? Or  
 917 might the quality of the auralization already be ac-  
 918 ceptable if it is perceived as "plausible", meaning the  
 919 simulation corresponds to a listener's expectation of  
 920 the corresponding real event [44] based on an internal  
 921 reference that is built up by everyday life expe-  
 922 riences [45]. In the context of virtual acoustic envi-  
 923 ronments, one could also say that authenticity means  
 924 that all perceivable "quality features" [46] of an acous-  
 925 tic environment are copied, while plausibility means  
 926 that only the features required for a specific purpose  
 927 are simulated. Following this definition, one could ar-  
 928 gue that it is enough for any application to strive for  
 929 "plausibility" as there is no point in reproducing un-  
 930 necessary features. This, however, implies that one  
 931 needs to know exactly which features are required for  
 932 a specific application. In the context of psychoacous-  
 933 tic experiments, that is not always possible, as the sole  
 934 purpose of such studies might lie in estimating which  
 935 features of a complex auditory scene are perceptually  
 936 relevant. This could be seen as an argument to always  
 937 strive for authenticity in the context of auralizations  
 938 for listening experiments.

939 From a more pragmatic point of view, we concluded  
 940 that, while there are some areas of virtual acoustics  
 941 where authenticity might be achievable, an authen-  
 942 tic auralization of complex acoustic scenarios such as  
 943 electric vehicle passages is very ambitious and comes  
 944 at the cost of high effort and low flexibility. Listen-  
 945 ing to the auralization results presented in Section 4,  
 946 it becomes clear that some small audible differences  
 947 would stand out in a direct A/B comparison, even  
 948 when the results are perceived as very similar to real-  
 949 life recordings. To design and fine-tune an AVAS sys-  
 950 tem for a specific existing vehicle, the overall auraliza-  
 951 tion should be as authentic or, when it comes to esti-  
 952 mating compliance with regulations, even as nume-  
 953 rically correct as possible. Nevertheless, for our goal  
 954 of investigating the human response to AVAS signals,  
 955 we argue that it is secondary whether the auraliza-  
 956 tion sounds exactly like an existing electric vehicle as  
 957 long as it is perceived as plausible and we are aware  
 958 of and have complete control over all signal proper-

ties. Even when the overall perception of a stimulus is  
 959 "only" plausible, there could nevertheless be some in-  
 960 dividual features that are perceived as authentic. For  
 961 example, in this work, we prioritize the AVAS sig-  
 962 nal auralization over the tire/road noise, which could  
 963 mean that the isolated AVAS signal is perceived as  
 964 authentic while the overall combination of AVAS and  
 965 tire/road noise is not.

966 Based on these considerations, we evaluated plau-  
 967 sibility in terms of "sounds like it could be an electric  
 968 vehicle passage" by performing a laboratory listen-  
 969 ing experiment with 20 participants. Additionally,  
 970 an indirect parametric comparison to the reference  
 971 recordings was performed by asking the subjects to  
 972 rate perceptual attributes such as annoyance and ve-  
 973 hicle speed for both auralizations and recordings. The  
 974 following sections describe the experiment setup and  
 975 discuss the implications of the obtained results.

### 977 5.1 Procedure and Stimuli

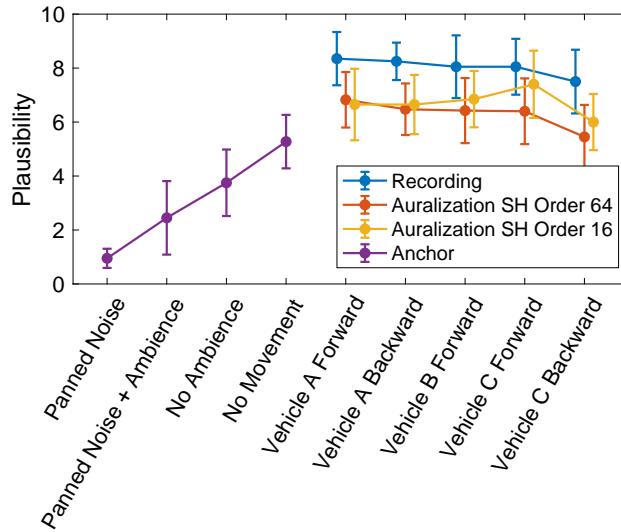
978 The listening experiment was divided into two parts:  
 979 In the first part, the participants were presented with  
 980 ten binaural in-situ recordings, two for each vehicle  
 981 and driving direction. The subjects were informed  
 982 that they were listening to real recordings and were  
 983 asked to rate perceived vehicle speed and annoyance  
 984 for each passage. When recruiting a group of non-  
 985 experts, one has to assume that not all subjects have  
 986 sufficient experience to rate plausibility without any  
 987 training phase, especially since electric vehicle sounds  
 988 are not yet well established in our everyday lives. The  
 989 purpose of this first experiment part was, therefore, to  
 990 familiarize the participants with the sound of electric  
 991 vehicle passages and, by this, build up an internal  
 992 reference while, at the same time, obtaining a "ground  
 993 truth" for the subjective vehicle speed and annoyance  
 994 ratings.

995 In the second part of the experiment, participants  
 996 were again presented with five out of the ten in-situ  
 997 recordings from the first experiment part, one for each  
 998 vehicle and driving direction. Additionally, the sub-  
 999 jects listened to 20 generated passages synthesized  
 1000 using the same vehicle velocity profiles as the cor-  
 1001 responding reference recordings. This means that,  
 1002 ideally, the generated signals in the second exper-  
 1003 iment part should be indistinguishable from the refer-  
 1004 ence recordings presented in the first experiment part.  
 1005 Of those 20 generated stimuli, ten passages were ren-  
 1006 dered with spherical harmonic order  $L = 64$ , five pas-  
 1007 sages were rendered with  $L = 16$ , and five passages  
 1008 were low-quality renderings included to act as an-  
 1009 chors. This group of low-quality renderings consisted  
 1010 of (i) amplitude-panned white noise, (ii) amplitude-  
 1011 panned white noise combined with a binaural ambi-  
 1012 ence recording, (iii) an auralization without ambience  
 1013 sounds, and (iv) auralizations without source move-  
 1014 ment for two different vehicles. For each stimulus, the

1015 participants were asked to rate perceived annoyance,  
 1016 vehicle speed, and plausibility compared to their internal reference based on the recordings presented in  
 1017 the first experiment.  
 1018

## 1019 5.2 Participants and Implementation

1020 The experiment was performed by 20 participants (12 male, 7 female, 1 preferred not to specify) recruited  
 1021 from Chalmers students and faculty members. The participants were aged between 22 and 37 years, with  
 1022 a median age of 27 years. All participants had self-  
 1023 reported normal hearing and an educational background in acoustics and gave their written consent  
 1024 for participation as well as collection and processing  
 1025 of their personal data. All stimuli were presented  
 1026 via calibrated headphones (Sennheiser HD 650), and  
 1027 the experiment was conducted using a HEAD acoustics SQala jury testing system. The order of stimuli  
 1028 within both experiment parts was randomized for  
 1029 each participant. All auralizations were rendered us-  
 1030 ing HRTFs measured for a HEAD acoustics HMS II.3  
 1031 artificial head [47], which has the exact dimensions  
 1032 and ear shape as the HMS V artificial head used for  
 1033 the reference measurements.  
 1037

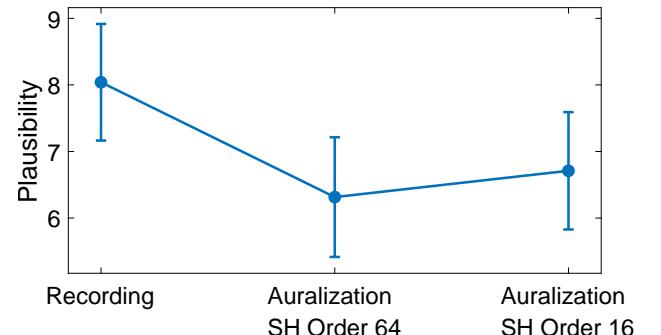


1038 Figure 12: Arithmetic mean of plausibility ratings  
 1039 with 95% confidence intervals. Vehicle B was only  
 1040 evaluated driving forward since its AVAS signal does  
 1041 not change with driving direction.

## 1038 5.3 Results

### 1039 5.3.1 Plausibility

1040 In the second part of the experiment, participants  
 1041 rated the plausibility of in-situ recordings, auraliza-  
 1042 tions with spherical harmonic order  $L = 64$  and  
 1043  $L = 16$  as well as low-quality anchor signals on  
 1044 a unipolar numerical 11-point interval scale ranging  
 1045 from the value 0 (“not at all plausible”) to the value



1044 Figure 13: Arithmetic mean of plausibility ratings  
 1045 with 95% confidence intervals, averaged over all ve-  
 1046 hicle types.

1047 10 (“extremely plausible”). Figure 12 shows the arith-  
 1048 metic mean and 95% confidence intervals of the ob-  
 1049 tained results. Independent of vehicle type, it can  
 1050 be seen that the in-situ recordings consistently scored  
 1051 the highest plausibility rating and that the anchor sig-  
 1052 nals are rated as least plausible, with the amplitude  
 1053 panned noise achieving the lowest plausibility score.  
 1054 Comparing the different vehicle types, the backward  
 1055 passages of vehicle C were rated as least plausible for  
 1056 both recordings and auralizations. This could be due  
 1057 to the fact that the vehicle C backward AVAS con-  
 1058 sists of a constantly repeating “plinging” sound that,  
 1059 in itself, could be perceived as artificial. Furthermore,  
 1060 there seems to be no consistent difference pattern be-  
 1061 tween both evaluated spherical harmonic orders.

1062 To further investigate the difference between in-situ  
 1063 recordings and auralizations, the plausibility ratings  
 1064 were averaged over vehicle type as shown in Figure 13.  
 1065 It can be seen that the recordings achieved a mean  
 1066 plausibility of around 8.0 out of 10 while the aurali-  
 1067 zation scored a lower average plausibility rating of  
 1068 6.3 for SH order  $L = 64$  and 6.7 for  $L = 16$ . The  
 1069 overall difference between those three stimuli groups  
 1070 was determined as statistically significant according  
 1071 to a repeated measures analysis of variance with  
 1072 Greenhouse-Geisser correction ( $F(1.502, 28.540) =$   
 1073 15.134,  $p < .001$ , partial  $\eta^2 = .443$ ). A Bonferroni-  
 1074 adjusted post-hoc analysis revealed a significant dif-  
 1075 ference between in-situ recordings and auralizations  
 1076 with SH order  $L = 64$  ( $MD = 1.725$ , 95%-CI[0.719,  
 1077 2.731],  $p < .001$ ) as well as between the recordings  
 1078 and auralizations with  $L = 16$  ( $MD = 1.330$ , 95%-  
 1079 CI[0.381, 2.279],  $p = .005$ ) but not between auraliza-  
 1080 tions with  $L = 16$  and  $L = 64$  ( $MD = -0.395$ , 95%-  
 1081 CI[-0.961, 0.171],  $p = .248$ ). This indicates that the  
 1082 auralizations are not perceived as plausible as the in-  
 1083 situ recordings and that there is no significant differ-  
 1084 ence in plausibility between spherical harmonic orders  
 1085  $L = 16$  and  $L = 64$  for this specific application. The  
 1086 fact that even the in-situ recordings were not rated  
 1087 as perfectly plausible shows that, despite the training  
 1088 phase in the first experiment part, not all participants

had a sufficiently strong internal reference to identify the authentic signals reliably. Even though not as good as the recordings, the auralization plausibility ratings are still relatively high on the scale and significantly better than for the low-quality anchor signals. Looking at the difference between the individual anchor signals, it becomes clear that both the added binaural ambient noise and the source movement are relevant features for the overall perceived plausibility since the signals rendered without those attributes are rated as less plausible than the complete auralizations.

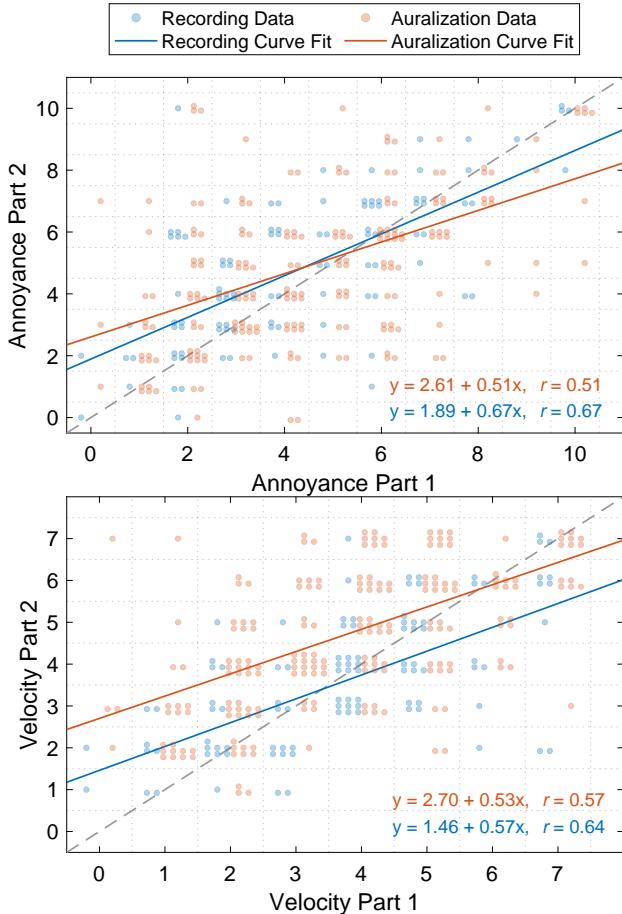


Figure 14: Annoyance and vehicle velocity ratings for both experiment parts and linear regression with correlation coefficient  $r$ . Observe that the used interval scale only allows for integer answers; data points with identical values were slightly offset to better visualize the distribution. The data combines the results for all evaluated vehicles; only auralizations with SH order  $L = 64$  are included.

### 5.3.2 Annoyance and Vehicle Velocity

In both experiment parts, the participant rated perceived annoyance and vehicle velocity for all stimuli. Thereby, annoyance was measured on a unipolar numerical 11-point scale labeled from 0 (“not at all

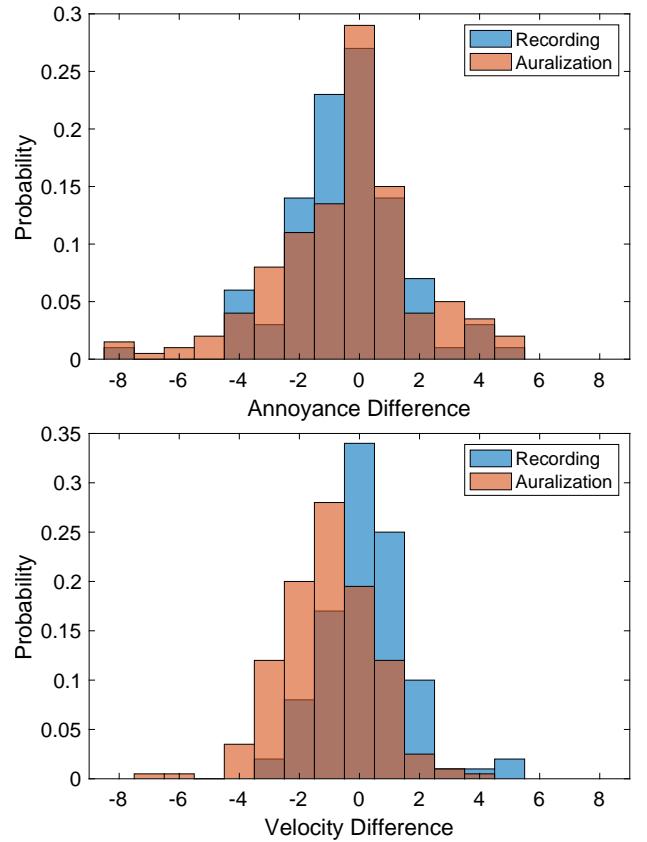


Figure 15: Distribution of differences in annoyance and vehicle velocity ratings between both experiment parts. The data combines the results for all evaluated vehicles; only auralizations with SH order  $L = 64$  are included.

annoying”) to 10 (“extremely annoying”) as recommended by ISO/TS 15666 [48], the perceived vehicle velocity was rated on an 8-point unipolar scale labeled as <5 km/h, 5-10 km/h, 10-15 km/h, 15-20 km/h, 20-25 km/h, 25-30 km/h, 30-35 km/h and 35-40 km/h. Since absolute velocity values are not of interest for the following evaluation, those eight velocity categories were translated to integer values ranging from 0 (<5 km/h) to 7 (35-40 km/h).

Using the ratings obtained for the in-situ recordings in the first experiment part as a reference allows for evaluating how well the auralizations presented in the second experiment part reproduce features relevant to the perception of annoyance and vehicle velocity. Since half of the in-situ recordings from the first experiment part were also repeated in the second part, we can additionally determine how consistent those subjective ratings are throughout the experiment. Figure 14 compares the annoyance and vehicle velocity ratings obtained for auralizations and recordings in the second experiment part to the results from the first part. Thereby, only ratings for stimuli pairs that exactly match each other were compared for each participant, i.e., values for the exact same recordings

in experiment parts one and two, as well as ratings for auralizations in part two, matched with ratings for the recordings in part one that they aimed to reproduce. If the ratings in experiment part two would perfectly match the results of the first experiment part, all data points in Figure 14 would lie on the identity line. However, that is not the case, neither for the recordings nor for the auralizations.

In order to better understand the difference between the results of both experiment parts, a simple linear regression was performed. The resulting regression lines presented in Figure 14 indicate that, for both repeated in-situ recordings and auralizations, participants tend to less extreme ratings in the second experiment part than in the first part, i.e., all regression lines have a similar slope smaller than one. This tendency could be statistically explained by a regression to the mean effect [49], i.e., assuming the annoyance and velocity ratings of each subject to be random variables with a certain distribution around a mean value, it is statistically more likely that subjects who gave an extreme rating in the first round tend to ratings closer to this mean value when repeating the experiment. Alternatively, this trend could be interpreted as a repetition priming effect [50], meaning that, after hearing all stimuli of the first experiment part, the participants might have adjusted their internal reference, resulting in more conservative ratings in the following part. Independent of the cause, this observation means that it is not sufficient for the perceptual validation of the auralization only to compare the difference between both experiment parts since even a “perfect” auralization that exactly reproduces the in-situ recordings would show this inconsistency in the subjective ratings.

While the correlation between the data for both experiment parts is slightly higher for the repeated in-situ recordings than for the auralizations, both data sets are only moderately correlated ( $r < 0.7$ ). This means that the relation of the data obtained for both experiment parts is not perfectly linear. Therefore, a simple linear regression might not be the most suitable tool for evaluating whether the difference between ratings for auralizations and in-situ recordings is statistically significant. Instead, the distribution of differences between both experiment parts was compared by subtracting the ratings obtained in the second experiment part from the ratings for the first part as shown in Figure 15. These distributions were then compared using a Wilcoxon signed-rank test, which showed that the distribution of differences in annoyance ratings compared to the reference in-situ recordings in the first experiment part, averaged over vehicles for each participant, does not significantly differ between recordings and auralizations ( $Z = -0.081, p = .936$ ). The difference in vehicle velocity rating was found to be statistically significant ( $Z = 3.825, p < .001$ ) which, in combination with the

shape of the distributions shown in Figure 15 and the offset between the recording and auralization velocity regression lines in Figure 14, leads to the conclusion that the auralization results in higher subjective vehicle velocity ratings than obtained for the corresponding in-situ recordings.

Based on the numerical comparison between auralizations and in-situ recordings discussed in Section 4, we assume that these differences in perceived vehicle velocity are related to the fact that the auralizations tend to show a more drastic change in time/frequency structure when the vehicle passes the observer which could be a consequence of inaccurate radiation directivities and missing environmental reflections. However, more research is needed to determine which features of an auralized vehicle passage influence the perceived velocity. While the perceptual validation results indicate that the implemented auralization method is unsuitable for experiments where the authenticity of the overall perceived vehicle speed is essential, it still allows for comparisons between stimuli, e.g., whether two different AVAS signals result in different speed perceptions.

## 6 Conclusion

This paper presented an auralization approach for electric vehicle passages based on in-situ measurements of three electric passenger cars. Different AVAS and tire/road synthesis methods were combined with radiation directivity and propagation models to generate pass-by auralizations suitable for AVAS-related psychoacoustic experiments. The numerical validation of the auralization results shows that, while the reproduction of the AVAS source signals is accurate, some discrepancies in the propagation modeling may be caused by inaccurate radiation directivities or missing environmental reflections. The auralization method achieved relatively high plausibility ratings in a perceptual evaluation, even though the generated stimuli were perceived as less plausible than in-situ recordings. While perceived annoyance ratings for the auralization results are consistent with the ratings for in-situ recordings, there is a statistically significant difference in velocity ratings between measurements and auralizations, which requires further investigation. Overall, we conclude that, while there are possibilities for improvement, the presented methods constitute a suitable foundation for AVAS-related listening experiments.

## Acknowledgments

This research was funded by FORMAS – a Swedish research council for sustainable development – under grant agreement FR-2020-01931. The HEAD-Genuit-

1239 Foundation provided the hardware used for the listening  
1240 experiment under grant P-22101-W.

## 1241 Data Availability Statement

1242 The data and audio examples associated with  
1243 this article are available on Zenodo under the  
1244 reference [doi.org/10.5281/zenodo.10610490](https://doi.org/10.5281/zenodo.10610490), all  
1245 relevant code is published on GitHub under  
1246 [github.com/leonpaulmueller/evat](https://github.com/leonpaulmueller/evat).

## 1247 Conflicts of Interest

1248 The authors declare no conflict of interest.

## 1249 Informed Consent Statement

1250 Informed consent was obtained from all subjects involved  
1251 in the listening experiment.

## 1252 References

- [1] IEA, "Global EV Outlook 2023," International Energy Agency, Paris, Tech. Rep., 2023.
- [2] Lykke Muller Iversen, Gerd Marbjerg, and Hans Bendtsen, "Noise from electric vehicles - 'State-of-the-art' literature survey," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, vol. 247, 2013, pp. 267–271.
- [3] EAA, "Environmental Noise in Europe," *European Environment Agency*, 2020.
- [4] Marlene Wessels, Sophie Kröling, and Daniel Oberfeld, "Audiovisual time-to-collision estimation for accelerating vehicles: The acoustic signature of electric vehicles impairs pedestrians' judgments," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 91, pp. 191–212, 2022, ISSN: 1369-8478. DOI: [10.1016/j.trf.2022.09.023](https://doi.org/10.1016/j.trf.2022.09.023).
- [5] UNECE, *Regulation No 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform provisions concerning the approval of Quiet Road Transport Vehicles with regard to their reduced audibility [2017/71]*, 2017.
- [6] Mendel Kleiner, Bengt-Inge Dalenbäck, and Peter Svensson, "Auralization - An Overview," *Journal of the Audio Engineering Society*, vol. 41, no. 11, 1993.
- [7] Michael J. Roan, Luke Neurauter, Miao Song, and Marty Miller, "Probability of detection of electric vehicles with and without added warning sounds," *The Journal of the Acoustical Society of America*, vol. 149, no. 1, pp. 599–611, 2021, ISSN: 0001-4966. DOI: [10.1121/10.0003386](https://doi.org/10.1121/10.0003386).
- [8] Marlene Wessels, Carolina Zähme, and Daniel Oberfeld, "Auditory Information Improves Time-to-collision Estimation for Accelerating Vehicles," *Current Psychology*, pp. 1–11, 2022, ISSN: 1046-1310. DOI: [10.1007/s12144-022-03375-6](https://doi.org/10.1007/s12144-022-03375-6).
- [9] Lisa Steinbach and M. Ercan Altinsoy, "Influence of an artificially produced stationary sound of electrically powered vehicles on the safety of visually impaired pedestrians," *Applied Acoustics*, vol. 165, p. 107290, 2020, ISSN: 0003-682X. DOI: [10.1016/j.apacoust.2020.107290](https://doi.org/10.1016/j.apacoust.2020.107290).
- [10] Nikolaos Kournoutos and Jordan Cheer, "Investigation of a directional warning sound system for electric vehicles based on structural vibration," *The Journal of the Acoustical Society of America*, vol. 148, no. 2, pp. 588–598, 2020, ISSN: 0001-4966. DOI: [10.1121/10.0001681](https://doi.org/10.1121/10.0001681).
- [11] Bernhard U. Seeber, Stefan Kerber, and Ervin R. Hafter, "A system to simulate and reproduce audio-visual environments for spatial hearing research," *Hearing Research*, vol. 260, no. 1-2, pp. 1–10, 2010, ISSN: 0378-5955. DOI: [10.1016/j.heares.2009.11.004](https://doi.org/10.1016/j.heares.2009.11.004).
- [12] Giso Grimm, Joanna Luberadzka, and Volker Hohmann, "A Toolbox for Rendering Virtual Acoustic Environments in the Context of Audiology," *Acta Acustica united with Acustica*, vol. 105, no. 3, pp. 566–578, 2019, ISSN: 1610-1928. DOI: [10.3813/aaa.919337](https://doi.org/10.3813/aaa.919337).
- [13] Institute for Hearing Technology - RWTH Aachen University, *Virtual Acoustics - A real-time auralization framework for scientific research*, Accessed on 2023-11-7. [Online]. Available: <http://www.virtualacoustics.org/>.
- [14] Daniel Gonzalez-Toledo, Luis Molina-Tanco, María Cuevas-Rodríguez, Piotr Majdak, and Arcadio Reyes-Lecuona, "The Binaural Rendering Toolbox. A Virtual Laboratory for Reproducible Research in Psychoacoustics," in *Forum Acusticum 2023 - 10th Convention of the European Acoustics Association*, 2023.
- [15] Ulf Sandberg and Jerzy A. Ejmont, *Tyre/road noise reference book*. Informex Ejsmont & Sandberg handelsbolag, 2002, ISBN: 9163126109.

- [16] EU, "Regulation (EU) 2020/740 of the European Parliament and of the Council on the Labelling of Tyres With Respect to Fuel Efficiency and Other Parameters," *Official Journal of the European Union*, 2020.
- [17] M. Strasberg, "Dimensional analysis of wind-screen noise," *The Journal of the Acoustical Society of America*, vol. 83, no. 2, pp. 544–548, 1988, ISSN: 0001-4966. DOI: [10.1121/1.396148](https://doi.org/10.1121/1.396148). [Online]. Available: <https://doi.org/10.1121/1.396148>.
- [18] Wolfgang Kropp, François-Xavier Bécot, and Stéphane Barrelet, "On the Sound Radiation from Tyres," *Acustica united with acta acustica*, vol. 86, pp. 769–779, 2000.
- [19] A.V. Oppenheim and R.W. Schafer, *Discrete-time Signal Processing* (Prentice-Hall signal processing series), 3rd. Pearson, 2014, ISBN: 9780132067096.
- [20] Beauchamp and W. James, "Additive Synthesis of Harmonic Musical Tones," *Journal of the Audio Engineering Society*, vol. 14, no. 4, pp. 332–342, 1966. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=1129>.
- [21] Julius O. Smith, *Spectral Audio Signal Processing*. 2011. [Online]. Available: <http://ccrma.stanford.edu/%5Ctextbackslashtextasciitildejos/sasp/>.
- [22] Couch, *Digital and Analog Communication Systems* (Prentice-Hall international editions), 8th ed. Pearson, 2013, ISBN: 9780132915380.
- [23] Martin A. Fischler and Robert C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981, ISSN: 0001-0782. DOI: [10.1145/358669.358692](https://doi.org/10.1145/358669.358692).
- [24] P.H.S. Torr and A. Zisserman, "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000, ISSN: 1077-3142. DOI: [10.1006/cviu.1999.0832](https://doi.org/10.1006/cviu.1999.0832).
- [25] Jan Jagla, Julien Maillard, and Nadine Martin, "Sample-based engine noise synthesis using an enhanced pitch-synchronous overlap-and-add method," *The Journal of the Acoustical Society of America*, vol. 132, no. 5, pp. 3098–3108, 2012, ISSN: 0001-4966. DOI: [10.1121/1.4754663](https://doi.org/10.1121/1.4754663).
- [26] Jens Forssén, Patrik Andersson, Penny Bergman, Krister Fredriksson, and Peter Zimmermann, "Auralisation of truck engine sound – preliminary results using a granular approach," *AIA-DAGA, Merano*, 2013. [Online]. Available: <https://research.chalmers.se/publication/188498>.
- [27] V. I. Lebedev and Dimitri N. Laikov, "A Quadrature Formula for the Sphere of the 131st Algebraic Order of Accuracy," *Doklady Mathematics*, vol. 59, pp. 477–481, 1999.
- [28] Earl G. Williams, *Fourier Acoustics*. London: Academic Press, 1999, ISBN: 9780127539607. DOI: [10.1016/b978-012753960-7/50006-1](https://doi.org/10.1016/b978-012753960-7/50006-1).
- [29] Gabriel Weinreich and Eric B Arnold, "Method for measuring acoustic radiation fields," *The Journal of the Acoustical Society of America*, vol. 68, no. 2, pp. 404–411, 1980, ISSN: 0001-4966. DOI: [10.1121/1.384751](https://doi.org/10.1121/1.384751).
- [30] Franz Zotter, "Analysis and Synthesis of Sound-Radiation with Spherical Arrays," Ph.D. dissertation, University of Music and Dramatic Arts, Graz, 2009.
- [31] Martin Pollow, "Directivity Patterns for Room Acoustical Measurements and Simulations," Ph.D. dissertation, RWTH Aachen, 2014.
- [32] Jens Ahrens and Stefan Bilbao, "Computation of Spherical Harmonic Representations of Source Directivity Based on the Finite-Distance Signature," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 83–92, 2020, ISSN: 2329-9290. DOI: [10.1109/taslp.2020.3037471](https://doi.org/10.1109/taslp.2020.3037471).
- [33] Jannik Theyssen, Thomas Deppisch, Astrid Pieringer, and Wolfgang Kropp, "On the efficient simulation of pass-by noise signals from railway wheels," *Journal of Sound and Vibration*, vol. 564, p. 117889, 2023, ISSN: 0022-460X. DOI: [10.1016/j.jsv.2023.117889](https://doi.org/10.1016/j.jsv.2023.117889).
- [34] Mansour Alkmim, Guillaume Vandernoot, Jacques Cuenca, Karl Janssens, Wim Desmet, and Laurent De Ryck, "Real-time sound synthesis of pass-by noise: comparison of spherical harmonics and time-varying filters," *Acta Acustica*, vol. 7, p. 37, 2023. DOI: [10.1051/aacus/2023029](https://doi.org/10.1051/aacus/2023029).
- [35] Dmitry N. Zotkin, Ramani Duraiswami, and Nail A. Gumerov, "Regularized HRTF fitting using spherical harmonics," in *Proc. IEEE Workshop Appl. Signal Process. Audio Acoust.*, 2009.
- [36] Matthias Frank and Manuel Brandner, "Perceptual Evaluation of Spatial Resolution in Directivity Patterns," in *Fortschritte der Akustik – DAGA*, 2019.

- 1438 [37] Tim Lübeck and Christoph Pörschmann, "Investigation of the minimum required spatial resolution of moving sound sources," in *Fortschritte der Akustik – DAGA*, 2023. 1439  
1440  
1441
- 1442 [38] Randall J. Allemand, "The Modal Assurance Criterion—Twenty Years of Use and Abuse," *Sound and Vibration*, vol. 37, pp. 14–23, 2003. 1443  
1444
- 1445 [39] Ilan Ben Hagai, Martin Pollow, Michael Vorländer, and Boaz Rafaely, "Acoustic centering of sources measured by surrounding spherical microphone arrays," *The Journal of the Acoustical Society of America*, vol. 130, no. 4, pp. 2003–2015, 2011, ISSN: 0001-4966. DOI: [10.1121/1.3624825](https://doi.org/10.1121/1.3624825). 1446  
1447  
1448  
1449  
1450  
1451
- 1452 [40] Jens Ahrens and Sascha Spors, "Reproduction of Moving Virtual Sound Sources with Special Attention to the Doppler Effect," in *124th Convention of the AES*, 2008. [Online]. Available: <http://www.aes.org/e-lib/browse.cfm?elib=14493>. 1453  
1454  
1455  
1456  
1457
- 1458 [41] "ISO 9613-1:1993 Acoustics — Attenuation of sound during propagation outdoors — Part 1: Calculation of the absorption of sound by the atmosphere," International Organization for Standardization, Tech. Rep., 1993. 1459  
1460  
1461  
1462
- 1463 [42] Jens Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. The MIT Press, 1996, ISBN: 9780262268684. DOI: [10.7551/mitpress/6391.001.0001](https://doi.org/10.7551/mitpress/6391.001.0001). [Online]. Available: <https://doi.org/10.7551/mitpress/6391.001.0001>. 1464  
1465  
1466  
1467  
1468
- 1469 [43] Fabian Brinkmann, Alexander Lindau, and Stefan Weinzierl, "On the authenticity of individual dynamic binaural synthesis," *The Journal of the Acoustical Society of America*, vol. 142, no. 4, pp. 1784–1795, 2017, ISSN: 0001-4966. DOI: [10.1121/1.5005606](https://doi.org/10.1121/1.5005606). 1470  
1471  
1472  
1473  
1474
- 1475 [44] Alexander Lindau and Stefan Weinzierl, "Assessing the plausibility of virtual acoustic environments," 2011, pp. 1187–1192, ISBN: 978-84-694-1520-7. 1476  
1477  
1478
- 1479 [45] Clemens Kuhn-Rahloff, "Realitätstreue, Natürlichkeit, Plausibilität, Perzeptive Beurteilungen in der Elektroakustik," 2012. DOI: [10.1007/978-3-642-22072-2](https://doi.org/10.1007/978-3-642-22072-2). 1480  
1481  
1482
- 1483 [46] Renato S. Pellegrini, "Quality Assessment of Auditory Virtual Environments." 1484
- 1485 [47] Christoph Pörschmann, Johannes M. Arend, and Raphael Gillioz, "Spherical Headgear HRIR Compilation of the Neumann KU100 and the Head acoustics HMS II.3 [Data set]," in *EAA Spatial Audio Signal Processing Symposium, Paris, France*, 2020. DOI: [10.5281/zenodo.3928465](https://doi.org/10.5281/zenodo.3928465). 1486  
1487  
1488  
1489  
1490  
1491
- [48] "ISO/TS 15666 Acoustics — Assessment of noise annoyance by means of social and socio-acoustic surveys," International Organization for Standardization, Tech. Rep., 2021. 1492  
1493  
1494  
1495
- [49] Adrian G Barnett, Jolieke C van der Pols, and Annette J Dobson, "Regression to the mean: what it is and how to deal with it," *International Journal of Epidemiology*, vol. 34, no. 1, pp. 215–220, 2005, ISSN: 0300-5771. DOI: [10.1093/ije/dyh299](https://doi.org/10.1093/ije/dyh299). 1496  
1497  
1498  
1499  
1500  
1501
- [50] Daniel L Schacter and Randy L Buckner, "Priming and the Brain," *Neuron*, vol. 20, no. 2, pp. 185–195, 1998, ISSN: 0896-6273. DOI: [10.1016/s0896-6273\(00\)80448-1](https://doi.org/10.1016/s0896-6273(00)80448-1). 1502  
1503  
1504  
1505

1506 **A Appendix**

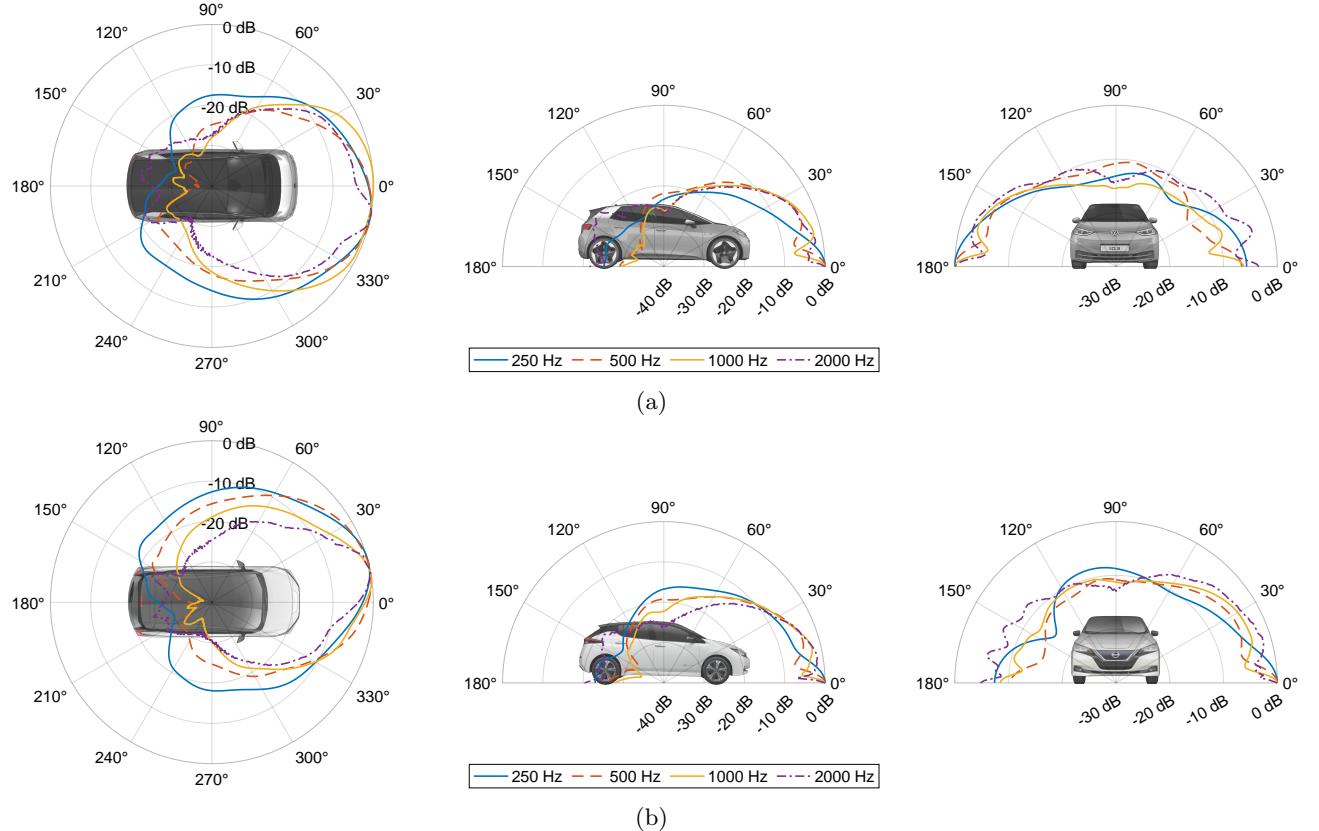


Figure 16: Vehicle B (a) and vehicle C (b) AVAS radiation directivity results of BEM calculation for the horizontal, median, and frontal plane. Normalized to the maximum for each frequency band. For vehicle A see [Figure 8](#)

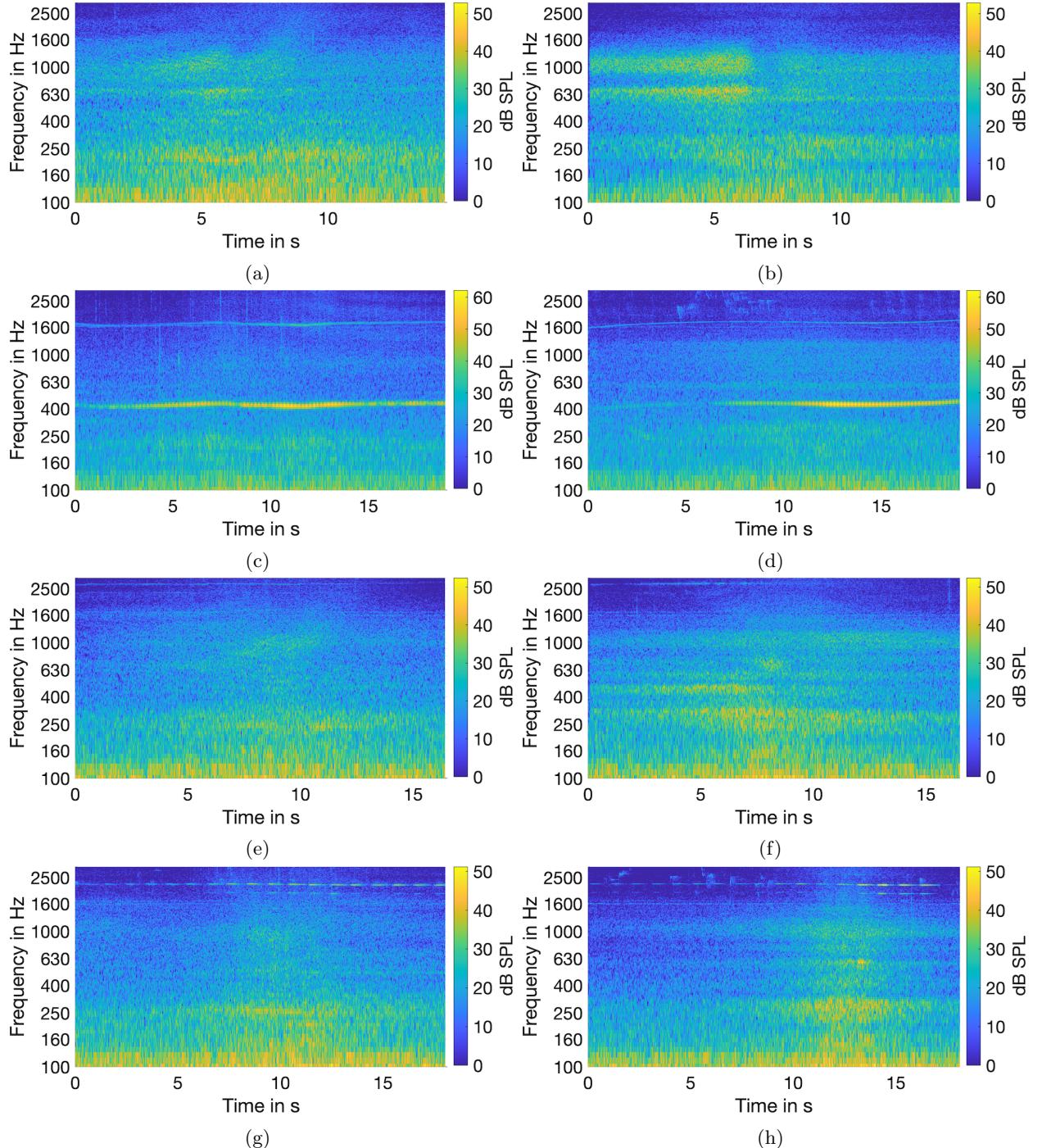


Figure 17: Vehicle A forward recording (a) and auralization (b), vehicle A backward recording (c) and auralization (d), vehicle C forward recording (g) and auralization (h) and vehicle C backward recording (i) and auralization (j). For vehicle B see [Figure 11](#).