



# Effects of rumble strip parameters on acoustics, vibration, and perception

Christian H. Kasess<sup>1,\*</sup> , Thomas Maly<sup>2</sup> , Piotr Majdak<sup>1</sup> , and Holger Waubke<sup>1</sup>

<sup>1</sup>Acoustics Research Institute, Austrian Academy of Sciences, 1040 Vienna, Austria

<sup>2</sup>Institute of Transportation, TU Wien, 1040 Vienna, Austria

Received 2 June 2021, Accepted 2 November 2021

**Abstract** – Rumble strips aim to alert the driver of dangerous situations via acoustic and tactile stimulation. They can, however, also lead to increased noise in the surroundings. Strip parameters and the vehicle type determines the size of these acoustic and vibratory effects. In our work, 16 rumble strip types (including strips with irregular spacing) were evaluated with respect to their effects on acoustics, vibration, and perception. Interior and exterior emissions caused by controlled pass-overs of a car and a truck were recorded. Annoyance, urgency, and reaction time of 16 listeners were tested in a laboratory experiment and models mapping acoustics to perception were developed. Steering wheel and seat vibrations were acquired. Our results show a high correlation of interior and exterior noise. The loudness level was the best predictor of perceived annoyance and urgency. Irregular spacing reduced the tonality at the cost of a difficult production process. While a large variety of strip types produced a sufficient acoustic effect in the passenger car, the acoustic response was lower in the truck. Combined with the low vibration levels found in the truck, an identification of an optimal strip type for both vehicle types turns out to be difficult.

**Keywords:** Rumble strips, Road traffic noise, Perception, Annoyance, Urgency

## 1 Introduction

Rumble strips are a sequence of grooves milled into the road surface or road markings raised on the road [1–4]. They can be placed along the boundaries of a road as longitudinal rumble strips or placed across a road as transversal rumble strips. Their aim is to alert the driver via acoustic and tactile stimulation, e.g., of the danger of running off the road or to slow down in front of a tunnel. The effects stem from the excitation of the tire when traversing the longitudinal height profile of the grooves or markings.

A considerable amount of various strip variants has been investigated with respect to the effect on the interior noise levels. The investigated parameters were the type (rolled grooves, milled grooves, raised markings), groove geometry (width, length, depth or height), as well as spacing between the grooves, see e.g. [1, 4]. Results from a regression model of the interior noise level change in a passenger car indicate that (1) the sound level increases with the speed but decreases with increasing spacing between the grooves, (2) milled rumble strips are louder than rolled ones, and (3) a concrete or wet surface reduces the effect of the strips [4]. When considering the vehicle type as parameter, for a passenger car or a pickup truck, a majority of the

strips produced sufficiently large sound level changes (with 4 dB as a threshold), but for larger vehicles such as trucks and busses, only special strip configurations (grooves wider than 30 cm and less than 90 cm apart) caused noticeable level changes [1].

When driving over the strips, the sudden onset of the sound may trigger a startle reflex of the driver [5], which may yield a dangerous situation. Thus, it is important to limit the strip effect [4]. Such a limitation may also be important from the perspective of residents living along the road [2, 6, 7] because the exterior noise resulting from the strip pass-overs may increase residents' annoyance. In an investigation of the exterior noise cause by three different groove shapes [2] (a rectangular cross-section, a cylindrical cross-section, and a sinusoidal shape [8]), the rectangular shape showed the largest SPL change, 2–5 dB higher than that caused by the cylindrical groove, which in turn was 1–2 dB larger than that caused by the sinusoidal pattern. In addition to the SPL change, other quantities may be relevant for the potential annoyance. For example, the grooves are usually periodically arranged [1, 4] yielding tonal components in the generated signal which are typically perceived as particularly annoying. Thus, in standardized models for environmental noise, tonality is often penalized [9, 10].

As a consequence, it is important to consider both, the interior and exterior noise. In [8], both effects were

\*Corresponding author: [christian.kasess@oeaw.ac.at](mailto:christian.kasess@oeaw.ac.at)

investigated for passenger cars and trucks, showing an exterior effect of the strips in the range of 5–7 dB and 3 dB (A-weighted SPL change), respectively. Inside the vehicles, the effects were different and depended on the type of the strips. For the car, the SPL changed by 13.3 dB and 14.1 dB for the sinusoidal and cylindrical strips, respectively. For the truck, the SPL changed by 2.6 dB and 7.6 dB, respectively, indicating a severe decline of the acoustic response for the sinusoidal strip in larger vehicles. The complex interaction between the strip type and vehicle was further confirmed in [11] showing that in the tested passenger car, sinusoidal profiles reduced the acoustic effects inside and outside, however, in the truck, the interior levels were barely noticeable. Considering all tested vehicle types, the cylindrical 300 mm groove seems to have provided a good compromise between the interior and exterior levels, both of which were correlated, in particular for the passenger car. In addition to the acoustic noise, vibration of the driver’s seat and steering wheel can be considered. For example, for sinusoidal profiles a large range of acceleration levels on the steering wheel and the seat was observed ranging from considerably lower to even higher values than those of cylindrical strips [8, 11] with the levels heavily depending on the strip design [11]. Furthermore, the type of tire has an impact on both the acoustics and the vibratory quantities [12].

Annoyance can be better described by quantities taking auditory perception into account such as loudness or sharpness, e.g. [13–15]. A study of the annoyance caused by passing over regular and irregular marking structures identified loudness and tonality as the main factors and a penalty of up to 10 dB in the A-weighted SPL was required to account for the tonality-based annoyance caused by the regular structure [16]. In this study, different scenarios were also generated using synthesized sounds based on the measurements. The temporal structure of a sound can also affect the warning effect in the interior of the vehicle. For example, a closer spacing of pulses leads to an increased urgency [17] as well as a decreased reaction time [3, 18]. Furthermore, temporal irregularity was also shown to capture the listeners attention. However, the relevant interval between the noise bursts (ranging from 100 ms to 300 ms) may be too large as that caused by typical rumble strips at typical vehicle speeds, (ranging from 10 ms to 20 ms). A further combination of the acoustic and the vibratory effects from rumble strips yielded mixed results [3].

Summarizing these findings, a thorough evaluation of the effects of rumble strips requires a combined examination of acoustical effects of both, interior and exterior noise considering perceptual-based measures combined with vibratory effects.

Thus, in this article, we investigate various effects of strip types aiming at reducing the noise burden but still keeping the alerting effect as high as possible. We considered various strip types, including sinusoidal strips and also irregularly spaced grooves. Especially for the irregular strip types, we expected a decreased tonality that may translate into a decreased annoyance and reaction time.

In our study, nine representative strip types were created, for which outdoor measurements and measurements in the vehicle interior were performed, considering two types of vehicles. Then, acoustical parameters were calculated and analyzed. Furthermore, additional types were synthesized, for which strips could not be manufactured in the first place. For such a large pool of conditions, we tested perceived annoyance, perceived urgency, and reaction time. The most promising candidates were manufactured and acoustically evaluated. Finally, we conclude with a general discussion.

## 2 Tracks and vehicles

### 2.1 Tracks

The test track was established on the hard shoulders of a highway near Vienna, Austria. The surface of the test track was asphalt. The top layer of the test track was of type SMA 11 S1 [19] and was established in 2011. At the time of measurement (2015–2017) the road was classified as status class 2 according to [20]. The surrounding terrain was dominated by arable land and no high growing crops such as corn were present at the time of measurements. The height of the shoulder varied roughly between 1 and 2 m along the test track. A service road about 3 m wide starting roughly 7 m away from the rumble strip was present. Nine strips were considered, see Figure 1. For better distinguishing the different variants, we denote the strip types used by the following scheme: first their approximate groove shape (a single capital letter), second the nominal groove spacing (for brevity given in decimeters), finally, if applicable, a type variation (lower-case letter). For example, C3r denotes a cylindrical groove with no or only very little edge rounding, the grooves are spaced 300 mm apart on average as the last letter “r” shows that the spacing is pseudo-random. Our motivation for the choice of the strips is explained in the next sections, where also details regarding the shape of the groove, the spacing, and their notation are described. Figure 1 shows a summary of the strips.

All rumble strips were milled and arranged in two lines in order to avoid large changes in the surrounding terrain and to meet the length restrictions of the hard shoulder. The inner line (with five rumble strips) was placed 30 cm from the outer edge of the road-marking of the outermost lane. The second line (with four rumble strips) was placed a further 56 cm next to the first line. Each strip type was 50 m in length and spaced approximately 30 m apart from each other. After milling the rumble strips, the parameters were validated with a custom-made depth measurement device based on a linear potentiometer for the depth and a rotary encoder for the distance measurement.

#### 2.1.1 Shape of the grooves

In the most common procedure, a cylindrical milling drum is lowered vertically resulting in a groove with a cross-section of a circular segment, e.g. [2, 4]. The patterns

label	C3*	C3r	C6	R3	R3i	R3o	S3c	S6c	S6
profile									
shape	cylinder	cylinder	cylinder	rounded	rounded	rounded	sine	sine	sine
pattern									
picture									
drum	flat	flat	flat	flat	flat	flat	curved	curved	flat
W / mm	300	300	300	300	300	300	100/300	150/300	300
L / mm	140	140	200	150	200/80	80/200	-	-	-
S / mm	300	300*	600	300	300	300	300	600	600
D / mm	10	10	10	7	12-5	5-12	6-16	9-16	0-7

**Figure 1.** Overview over measured strip types.  $W$  denotes the width,  $L$  the length and  $D$  the depth of the grooves.  $S$  denotes the spacing. The strip marked with \* was only measured in the inside of the vehicle. + denotes the average spacing for irregularly spaced grooves.

can vary in spacing between the grooves, the length, the depth, as well as the width of the groove (perpendicular to the driving direction), all of which are known to influence the noise emissions outside (see e.g., [11]) as well as inside the vehicle [4, 11]. With no or almost no rounding at the groove edges, we refer to this shape as *cylinder shape* and denote the corresponding strip types with the letter “C”. Being a very common strip type, our motivation for including this shape is to provide a widely used reference. Note that such cylindrical patterns are sometimes also referred to as rectangular due to the shape of the patches when viewed from above, e.g. [11]. Note, however, that there are grooves with a truly rectangular depth profile [2], but they were not considered in this study as they are difficult to produce.

In contrast to the cylindrical designs, sinusoidal strips do not have grooves with a flat section in between. Rather, the depth of the strip changes continuously throughout. The main idea is that the tire follows the depth profile more smoothly resulting in reduced impulsive excitation of the tire and thus reduced noise [2]. Despite low acoustic emissions, the sinusoidal shapes can produce high levels of vibration

on the steering wheel, rendering them potentially good candidates for silent but still effective strip types (e.g. [8, 11]), this being our main motivation in including this type in our investigation. We denote all *sinusoidal* strip types with the letter “S”, but we differentiate between two variants of sinusoidal strips. First, a cylindrical milling drum can be used with sinusoidal vertical lowering resulting in a constant depth in the transverse direction. Alternatively, a *curved* milling drum can be used resulting in a varying depth and width of the strip in the transverse direction. We highlight the latter strip types by the letter “c” in the postfix.

As an intermediate step, the groove edges of cylindrical designs can also be *rounded*, resulting in a more smooth shape but still a clearly discernible flat section between the grooves. We denote the corresponding strip types with the letter “R”. The main motivation of this strip type is to allow the tire to better follow the depth profile and thus reduce excitation similar to the sinusoidal strip, albeit to a lesser degree. Note, however, that distinguishing cylindrical and rounded variants is only a rough categorization as there is typically also a small degree of rounding for the cylindrical grooves due to the production process.

**Table 1.** Psychoacoustic parameters determined from the measurements.

Quantities	Unit	Method	Description
$L_{N5}, L_{N50}$	phon	DIN45631/A1 [22], free field	loudness level
$S_{D5}, S_{D50}$	acum	DIN45692 [23], free field	sharpness
$S_{A5}, S_{A50}$	acum	[24], free field	sharpness
$R_5, R_{50}$	asper	[25]	roughness
$T_5, T_{50}$	tu	[24, 26]	tonality

Typically, the groove shape is symmetrical along the transverse direction. By tilting the mill to a side, slanted grooves can be easily produced, i.e. grooves either deeper towards the traffic lane or towards the road shoulder. We tested the *inward* and *outward* slanted grooves to test whether this changes the directivity of the noise emission, and thus reduces the outdoor annoyance without affecting the required interior effects. They were tested for the rounded shape only and are denoted with the letter “i” and “o”, respectively, in the postfix.

### 2.1.2 Spacing of the grooves

In general, grooves of rumble strips are regularly spaced resulting in a tonal noise such that at constant speed of the vehicle the produced fundamental frequency remains constant. Tonality is a particularly annoying component of noise and typically leads to penalty terms in standards or regulations [9, 10]. Thus, regularly spaced grooves may be especially annoying for the residents. A central part of this study was to investigate whether irregular spacings are able to reduce outdoor annoyance.

To this end, we investigated grooves with pseudo-random spacing. This condition is denoted by the letter “r” in the postfix. The intervals were generated using a pseudo-random number generator (implemented as *runif* in the package *stats* from R [21]). In order to obtain an average pseudo-random spacing for 300 mm, the lower limit of the spacing was set to 200 mm and the upper limit to 470 mm. Then, a uniformly distributed random *inverse* spacing was generated having a mean distance of 300 mm. The rationale behind this approach was to generate a flat noise spectrum at the same mean excitation frequency as the regular grooves. Note that irregular groove spacing cannot be produced in the typical procedure of periodically lowering the milling drum while moving forward. For our study, every single groove was created by placing the milling drum above the corresponding position and then lower it to the appropriate depth without any longitudinal motion during the milling process – an approach not achievable on a large scale. Note that by doing so, truly cylindrical grooves were produced, i.e., there was no rounding of the edges whatsoever.

## 2.2 Vehicles

For the pass-overs, we used a passenger car and a truck. The passenger car was a BMW 320d (2003) sedan with an axial spacing of 2.725 m and “Speed Grip 2” (Semperit AG, Austria) tires of dimension 205/55 R16H. The truck was an

MAN 84S (2014) and provided by the Austrian highway authority ASFINAG. Front tires were “Ultra Grip WTS” (Goodyear Corporate, Ohio, USA) of dimension 385/65R22.5. In the period between the first and second measurement campaign the front tires had to be replaced with a “Ultra Grip max. S” (Goodyear) of the same dimension. The second axle (3.4 m spacing to the front axle) had mounted “Ultra Grip WTD” (Goodyear) of dimension 315/80R22.5 on twin wheels. The third axle (1.4 m apart from the second axle) was equipped with “Ultra Grip WTS” (Goodyear) of dimension 385/65R22.5.

All measurements for the passenger car were made with 100 km/h and all measurements for the truck were performed at a speed of 80 km/h.

## 3 Acoustics

We were interested in the effect of rumble strips on various acoustic descriptors, from the perspective of the residents and the driver. To this end, the recorded signals considered outdoor and interior settings. Based on those recordings, acoustic parameters, summarized in Table 1 were calculated and analyzed.

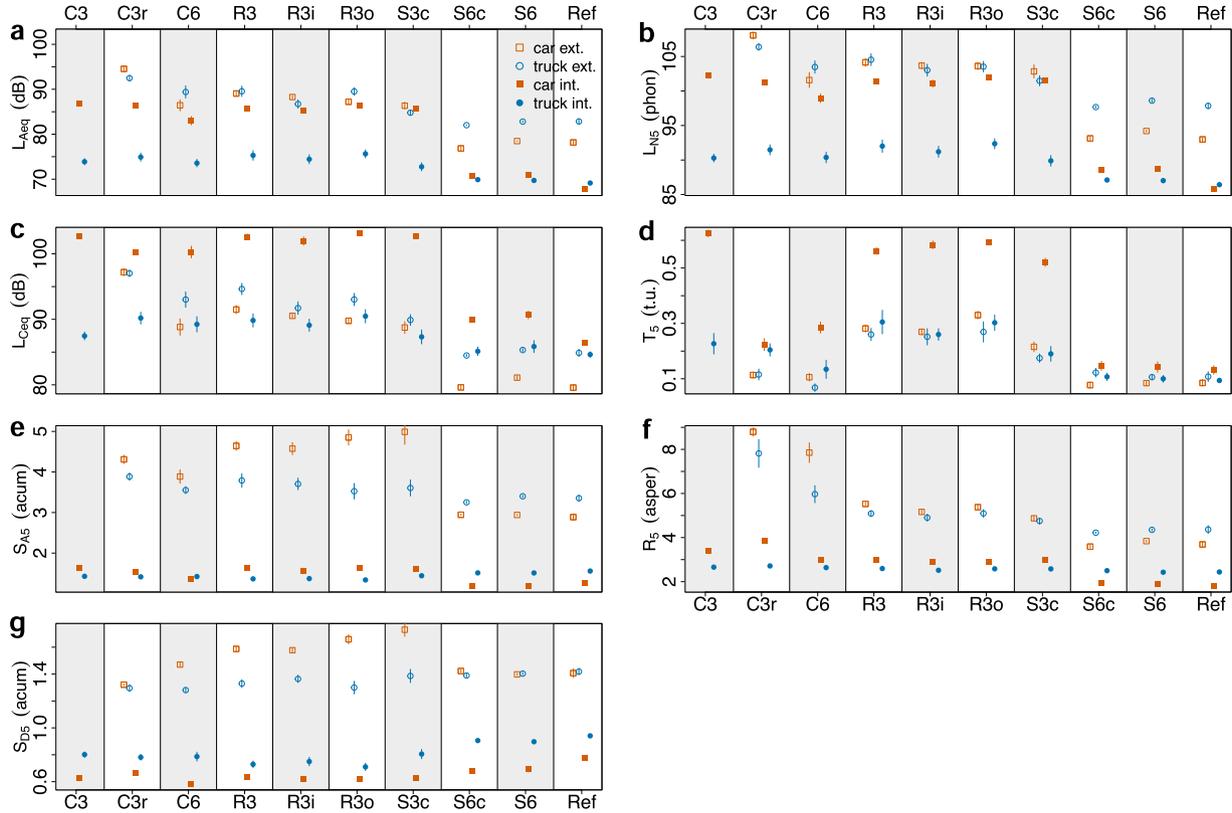
### 3.1 Methods

#### 3.1.1 Outdoor

Two microphones were placed in a distance of 7.5 m and 25 m of the strip edge, both 1.2 m above the road surface. The microphones were placed at the center of the 50 m long strip and their distances were adjusted according to the line of rumble strips under investigation. The microphones were 1/2” TEDS microphones (G.R.A.S. 46AE) connected to a 24-channel ICP-module (DIC 24, HEAD Acoustics). Additionally, a head-and-torso simulator (HATS HMS IV, HEAD Acoustics) was placed next to the 25 m microphone in order to obtain binaural recordings. In the acoustical analyses, the channel with the higher value of the respective acoustic feature of the binaural channels was used.

A light barrier was placed in a distance of 3 m from the edge of a strip. It was connected to the data acquisition system and together with two reflectors placed on the side of each vehicle was used to determine the timing of the vehicle entering a strip and the vehicle’s speed.

In order to determine the coverage of the tyre on the strip, a camera (GoPro Hero 3+ for the passenger car, GoPro Hero 4 for the truck) was attached to each vehicle, recording the front tire position during each pass-over.



**Figure 2.** Acoustic descriptors (see Tab. 1) for exterior (microphone at 7.5 m, open symbols) and interior (HATS, closed symbols) recordings, and both vehicle types (encoded by the color). Shown are a) the A-weighted equivalent sound pressure level  $L_{Aeq}$ , b) the peak loudness level  $L_{N5}$ , c) the  $L_{Ceq}$ , d) the peak tonality  $T_5$ , e) the peak Aures sharpness  $S_{A5}$ , f) the peak roughness  $R_5$ , and g) the peak DIN-sharpness  $S_{D5}$ . The error bars show the 95% confidence intervals.

Each cross-section, i.e., two rumble strips, was acquired on a single day. For the passenger cars, at least 10 and up to 24 (median 16) pass-overs were measured. For the truck, 10–19 pass-overs (median 14.5) were recorded. Note that the vehicles had to run a loop of roughly 14 km from one highway exit back to the next and thus the slightly lower number of recordings for the truck is a consequence of its lower speed. In addition, eight recordings (for the passenger car, and seven recordings for the truck) of regular pass-bys next to the rumble strips were recorded in order to obtain a reference condition without any strip effects.

### 3.1.2 Interior

For the interior measurements, the HATS was placed inside the vehicle. In the passenger car it was placed on the co-driver’s seat and an additional microphone (G.R.A.S. 46AE) was placed between the HATS and the driver. In the truck, the HATS was placed between the driver and the co-driver. The additional microphone was placed between the driver and the HATS, roughly at the height of the driver’s ears.

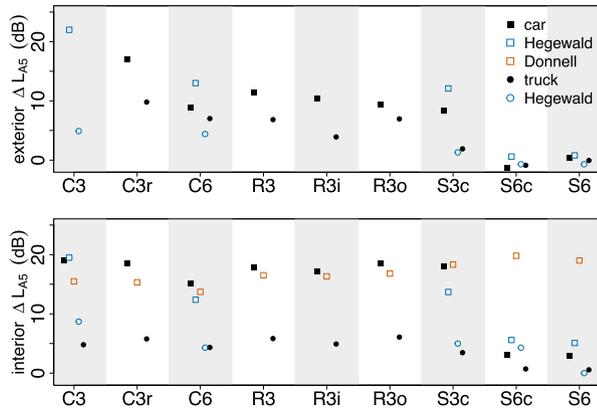
Similar to the outdoor measurements, the camera was used to monitor the tire coverage. Further, the vehicle’s speed was monitored by means of GPS-logging and the radio was turned off.

Measurement data were acquired for each vehicle and campaign in a single day, however, truck and passenger car measurements had to be done on separate days. Overall, for the passenger car 18–21 (median 21) pass-overs were recorded, and for the truck 15–18 (median 18) recordings were performed.

### 3.2 Data analysis

The recordings were pre-selected on the basis of the tire coverage of each pass-over. From the camera recordings, the coverage rating was determined at the beginning, the middle, and the end of each pass-over. The coverage rating was manually judged by visual inspection and resulted in a zero for the perfect coverage and  $\pm 5$  for a complete miss (tire outside of the groove). Values from 1 to 4 implied an approximate coverage of 90, 75, 50, and 25%, respectively. For the exterior recordings, in order to penalize poor tire coverage more heavily, the root-mean-square value of the coverage at the beginning, middle, and end of the pass-over was calculated and the pass-overs were ranked accordingly. For example, coverages of (0,1,1) at beginning, middle, and end and (0,2,0) would have the same arithmetic mean of 2/3 but a different RMS.

Figure 2 shows the acoustic descriptors calculated for both vehicle types and for the 7.5 m exterior and



**Figure 3.** Change of sound level  $L_{A5}$  relative to the reference levels. Top: Exterior. Bottom: Interior. Black filled symbols:  $L_{A5}$  differences determined from the measured data. Orange symbols: Predictions obtained for the difference between peak level and ambient level based on the model from [4] using our groove parameters. Blue symbols: Peak level differences as obtained in [11]. Note that for the latter the rumble strip design closest to the published ones was chosen.

HATS-interior recordings as a function of the strip type, ordered as in Figure 1. Only pass-overs with the coverage grade of 3 or better were considered. The error bars show the 95% confidence intervals from a t-statistic.

Additionally, relative sound levels  $L_{A5}$ , with respect to the  $L_{A5}$  of the reference recordings, i.e., levels obtained without any strips, are shown in Figure 3. This figure also includes data from [11] for the rumble strips with parameters most similar to the ones denoted and predictions for the interior noise calculated with the regression model from [4] using the strip parameters from our current study.

### 3.3 Results and discussion

The SPL and loudness for both, exterior and interior, (Figs. 2a and 2b) decreased from C3 to C6 and S3c to S6c, i.e., with the increase of the groove spacing. The sinusoidal strip types resulted in the lowest levels. Especially, the sinusoidal strips with 600 mm groove distance, S6c and S6, yielded nearly the same pass-by levels as no strips at all. The sinusoidal strip with 300 mm groove distance, S3c, showed the next larger sound pressure/loudness. The rounded variants lie in between with all variants including inward and outward slanted grooves (R3i and R3o) exhibited quite similar descriptors, thus showing no substantial effect of the groove asymmetry. Thus, as expected, the data indicate a general trend to decreased sound pressure and loudness levels for increased groove spacings and rounder edges.

For the exterior recordings (open symbols in Fig. 2) car and truck produced similar noise levels for most rumble strips, but due to the higher reference pass-by noise, the differences are less pronounced for the truck. As S6 and S6c produced low emissions, for these strips, the truck's pass-by levels were larger than those of the car. The pseudo-random strip C3r was by far the loudest strip.

Similar to the exterior case, for the interior recordings (filled symbols in Fig. 2) the difference between the reference background noise level and the rumble strip pass-over was reduced for the truck. In the interior case, however, this is caused by a reduced response to the rumble strip as background levels are comparable. The much lower differences for the truck may be a consequence of the vehicle itself, probably due to a much higher insulation of the driver cabin and a larger damping from the tire to the seat and steering wheel.

The general trends are in agreement with those from Table 13 in [11] (blue empty symbols in Fig. 3) up to differences of 5 dB in the respective conditions. Note, however, the comparison might be confounded for some reasons. First, the parameters of the strip designs are not exactly the same. Second, the vehicles and tires are different across the two studies, which may lead to differences in the reference noises. Third, for the exterior recordings, the speed of the passenger car was 80 km/h which was lower than in our study. Fourth, in [11] the  $L_{AFmax}$  was used. Finally, exterior recordings were performed closer (5 m instead of 7.5 m) resulting in potentially different reference levels due to e.g. different contributions of engine and rolling noise. Especially for the truck the lower differences for the rumble strips in the exterior noise and the high difference of the pass-by noise between truck and passenger car of almost 11 dB may indicate a potentially higher pass-by noise of the truck in [11].

The interior levels for the car were in addition compared to predictions of [4], which is a noise model comprising a linear regression with independent variables speed, groove length, width, depth, and spacing. For all strips, we used asphalt as the pavement type, dry condition, milled method of producing the rumble strips, the outer shoulder as the location of the strip, and  $0^\circ$  as the angle of incidence. The orange empty symbols in the lower panel of Figure 3 show these predictions and the values indicate a good general agreement of the model with the measurements. Compared to the measurements, the model underestimates the effect by 1–4 dB. For the sinusoidal strips with 600 mm spacing (S6 and S6c), however, there is a severe overestimation of the effect which is simply due to the fact that the model was not for this type of strip as no sinusoidal strips were measured in [4]. More specifically, the actual longitudinal groove profile was not part of the model at all which may explain the differences in model error between the rounded and cylindrical patterns.

Concerning the tonality (Fig. 2d), for the pseudo-random strip C3r, it was lower than that for comparable regular strips, e.g., C3, R3, and S3c. This confirms our hypothesis, that the pseudo-random groove distance reduces the tonality. Note that for 600 mm groove distance, the tonality was low even for regular groove distances, which may be partly because of the low fundamental frequency of 46 Hz or 37 Hz for the car and the truck, respectively, for this large groove distance. Thus, at a later stage of our study, we investigated the pseudo-random strip with an average groove distance of 600 mm to look at this factor in more detail (see Sect. 5).

The roughness and sharpness (Figs. 2e–2g) increased in the exterior recordings for most strip types, when compared to the reference. The roughness increased substantially for the non-rounded grooves C3r and C6, especially for the passenger car. There was a large difference between C6 and R3, and the effect can be attributed to both parameter changes between these two strip types: groove spacing and profile. In order to disentangle the contribution of these two factors, at a later stage of our study, we investigated recordings of additional strip types, more systematically varying the groove distance (see Sect. 5). For the interior recordings the differences in roughness and sharpness between rumble strips and the reference were far less pronounced and not much conclusive. Concerning the factor vehicle type, for the truck, the strip type had a smaller effect on the roughness and sharpness than for the car, in both, exterior and interior recordings.

## 4 Perception

In this section, we tested the effect of the strip types on the perceived annoyance, urgency, and the reaction time in human listeners. The aim was to relate the findings based on acoustic descriptors to the perceptual perspective of the driver and residents.

### 4.1 Methods

#### 4.1.1 Listeners and equipment

Sixteen listeners (9 female, average age  $26 \pm 5.5$  years) participated in the study. Audiograms were measured at frequencies from 0.125 to 8 kHz and none of the listeners showed hearing loss. The tests were carried out in a sound-proof chamber with a background A-weighted SPL of 17 dB. The test protocol was implemented in a custom-made program on a standard personal computer. Keyboard and mouse were used for listener's input. Tests including the audiogram (for most listeners this was already available) were performed in a single session which lasted typically around 4–5 h. Listeners could take breaks between the tasks as well as between the runs of the different tasks as outlined below.

Playback of the stimuli was done using a computer-controlled digital equalizer (PEQ V, HEAD acoustics) connected to the PC via USB 2.0 and driving open dynamic headphones (HD 650, Sennheiser). The sampling rate was 48 kHz.

#### 4.1.2 Stimuli

Binaural HATS recordings of the pass-overs from Section 3.1 were used. These recordings took place on a highway during regular operation and thus, only a small number of recordings without traffic noise for only a limited number of strip types was available. As a consequence, an important aspect of this work was to extend the set of rumble strips to be evaluated by using synthetic data of designs that could not be produced on the test track. To this end,

the data was augmented by simulating additional strips types. The simulations were based on individual single-groove signals (derived from single groove recordings or inverse filtering), considered BEM-based noise propagation model, moving sources, various groove parameters, high-frequency noise, and individual level calibration. Noise propagation and moving sources were only needed for the exterior noise simulation. More technical details on the simulations of outdoor and interior measurements can be found in the Appendix of Sections A.1 and A.2, respectively.

In total, for each vehicle type, interior and exterior measurements of 21 and 18 strips were simulated, respectively. As a first step, we began the simulations of the measured C3 (interior only), C3r, C6 (interior only), R3, and S3c. These simulations allowed us to informally verify the simulated results.

Based on the sound of a single groove similar to those of C3r, for the exterior we simulated R3r, R4, R4r, R6, and R6r by changing the distance between the simulated grooves. This was aimed at a deeper investigation of the effects of groove spacing and its regularity. In order to investigate these effects in sinusoidal strip shapes, we simulated S3 and S4 based on the simulations of S3c. For the interior noise this was done in a similar fashion except that the single groove signal was derived from an inverse filtering of the recordings of C3r since the signal-to-noise ratio in particular in the truck was relatively poor such that the single groove pass-over could not be used.

In addition, the random strip R3r was simulated with two additional variations: with a 4-grooves period (R3r4) and with a 20-grooves period (R3r20). Also, the random strip R4r was additionally simulated with a 3-grooves period (R4r3) and with a 15-grooves period (R4r15, exterior only). The reasoning for these semi-periodic strips was to produce a pattern which can be more easily produced on the road than the fully random strips, with the downside of producing a potentially perceivable low-frequency modulation.

Finally, to further increase the variety of signal classes, the effect of C3r (interior only), C6 (interior only), R3, R4, R3r, and R6 was additionally simulated with amplitudes scaled by  $-6$  dB for the passenger car and  $+4$  dB for the truck before adding the regular vehicle noise in the simulation.

#### 4.1.3 Annoyance rating

Listeners received written instructions containing descriptions of the perceived annoyance and the procedure. They were asked to base their ratings on imagining how annoying and distracting they would perceive the noise, if they were subjected to it on a regular basis [27, 28]. Following the instructions, listeners performed a training covering a wide range of stimuli. After the training, listeners were allowed to adapt their rating range in case they felt uncomfortable with their initial choice. At this point listeners also had the opportunity to clarify open issues.

The annoyance ratings were collected using a free magnitude estimation of the perceived annoyance [14, 29].

At the begin of a trial, a stimulus was played with the playback indicated visually. The listener was then asked to input a numerical rating corresponding to the perceived annoyance. While listeners were free in choosing their starting value, they were instructed to avoid extremely high or low starting values in order to stay within a comfortable range of numbers. Listeners were asked to perform a proportional rating, i.e. doubling the annoyance being reflected in doubled rating. They were also explicitly told to keep their rating scale constant within and across all runs. During the playback the input field was disabled to avoid a premature judgement. After having provided the rating, listeners continued by pressing a button. Each listener was tested in four runs with a break of at least 5 min between each run. The total experimental time for all 464 items was about 60 min, depending heavily on the listener as the experiment was self paced.

Whenever a listener reported a typo (13 in total) that could be uniquely identified (11 in total), either the rating was corrected when simple to reconstruct from the listener's comments, or the rating was discarded (2 in total). Furthermore, two times listeners entered the value 0 which were also ignored. Thus, a total of 4 inputs out of 7424 were discarded. The consistency across runs was analyzed using the correlations of the average log-ratings per condition across runs. The six correlation coefficients were mostly above 0.8 and always above 0.6.

For the analysis, the base-2 logarithm was applied to the ratings in order to reflect a doubling of the perceived annoyance in an increment of the log-rating. Then, for each condition, a histogram of all log-ratings was calculated and individual log-ratings outside of the threefold standard deviation were discarded as outliers. Listener-specific log-ratings were calculated as the median log-rating across runs per listener and condition. To normalize across the listeners, the total mean of the log-ratings per listener was removed from the listener-specific log-ratings [14, 29]. Finally, the group annoyance ratings were calculated by arithmetically averaging the listener-specific log-ratings across listeners.

#### 4.1.4 Perceived urgency

The stimuli were a subset consisting of the beginning of a pass-over because that part of a pass-over best represented the transient characteristics of the recordings. The stimuli were cut such that after 0.7 s of normal interior noise the rumble strip noise started for 1.8 s. Listeners received written instructions and performed a training run before the test.

The listeners were asked to rate the change of the interior noise to the rumble strip. The listeners rated the urgency on a fixed scaled consisting of five main levels labeled: not at all, little, medium, very, extremely urgent. Listeners had also the possibility to provide ratings between the level, such that a total of nine levels was available. The fixed scale, contrary to annoyance, was chosen in order to obtain a clear fixed rating for the case of no urgency at all. As catch trials, regular interior noise samples without

a rumble strip were included in the tests. Four runs were done with each condition being present twice in each run. The urgency task was performed after the annoyance rating task, and the total experimental time for all 432 items was also about 60 min, again depending heavily on the listener as the experiment was self paced.

For the analysis, the mean rating was calculated per condition and subject. The consistency across runs was analyzed similar to that for the annoyance. All correlation coefficients were above 0.7, with the listener-specific median across condition between 0.8 and 0.95.

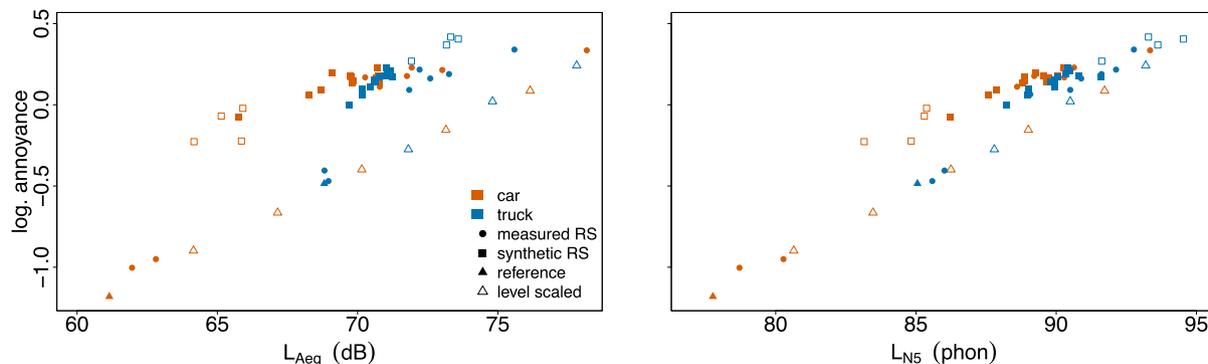
#### 4.1.5 Reaction time

A continuous background noise was generated by concatenating recordings with cross fading. Rumble strip noises of 1.2 s duration were cross-faded (in the case of measured recordings) or overlaid (in the case of simulations). The duration between the strip-noise events was random in the range between 3 s and 7 s. Catch trials were considered by cross-fading to the background noise, without the strip noise.

The test was arranged in six runs for the passenger car and six runs for the truck. Each run consisted of 54 strip noises (each condition twice) including catch trials. Note that the long-period semi-periodic strips R3r20 and R4r15 were too long to evoke a difference in the reaction time when compared to the pseudo-random strip R3r, thus were not tested. A total of 10,368 events were presented.

During the presentation listeners had to perform a distracting task in which a randomly moving target on a screen had to be followed by a cursor controlled by the computer mouse [18]. In order to increase the cognitive load, listeners had to use the hand with which they usually do *not* use the mouse. With the other hand, whenever they perceived a signal change, the listener had to press the space bar as quickly as possible. To ensure a proper recording of the reaction times, a custom-made hardware recorded a pulse on a cross-fade as well as a pulse on a space-bar hit. The reaction time task was performed last and took about 120 min for the 648 items.

For the analysis, 17 responses were excluded due to an unreasonably short reaction of less than 0.1 s. Furthermore, 198 reactions occurred after 1.2 s, i.e. after the end of the rumble strip noise, and were counted as not reacted. In total, 775 events lead to no reaction at all. Outlier detection (as that for the annoyance test) detected 16 outliers which were discarded. If a listener reacted only in less than half of all trials within a single condition, the corresponding condition was classified as "not detected" by that listener and excluded from the averaging across listeners. In total, 15 listener-specific conditions were excluded. Consistency between the first and second half of the experiment was also checked as for annoyance and urgency, however, because of so many factors playing a role in a reaction time test, the correlation coefficients were rather low, ranging from 0.2 to 0.8 for the car, and 0.1 to 0.7 for the truck.



**Figure 4.** Log-ratings of the perceived annoyance as a function of A-weighted SPL,  $L_{Aeq}$  (left panel), and peak loudness level,  $L_{N5}$  (right panel). Each symbol represents an average across all listeners. Squares: rumble strips. Triangles: reference without rumble strips. Open symbols: Conditions scaled in the level. Filled symbols: actually measured or simulated conditions. The color denotes the vehicle type.

## 4.2 Results

### 4.2.1 Perceived annoyance

The left panel in Figure 4 shows the log-ratings as a function of the A-weighted SPL,  $L_{Aeq}$ . Each symbol denotes the average across listeners for a single condition. Color decodes passenger car (orange) and the truck (blue). Also, the log-rating of the reference condition, i.e., regular pass-bys without any rumble strips are shown. Conditions obtained by scaling the level are shown as open symbols.

The references were differently perceived than the rumble strip pass-overs. The relation between level and annoyance seems to differ between the reference and rumble strip conditions, although at high levels the two different condition sets seem to get close to each other. The only exceptions to this difference are the 600 mm sinusoidal strips, S6c and S6, which were at the level of the regular pass-bys, indicating that these strips did not evoke any excess annoyance. Concerning the periodic pseudo-random strips there did not seem to be any difference from the purely random strips of the same period.

The analysis of the log-ratings as a function of peak loudness level,  $L_{N5}$  (Fig. 4, right panel) does not add much more insights, showing that based on any of these two single acoustic descriptors, the perceived annoyance as an effect of the rumble strip cannot be accurately predicted. Thus, similarly to [13] a step-wise scheme was used to determine the optimal model considering all the acoustic descriptors from Table 1. The Bayes information criterion (BIC) [30] was used to judge the trade-off between model accuracy and complexity. The function *stepAIC* [21, 31] was used to perform this step-wise selection applying a forward-backward scheme, i.e., variables may be discarded and re-included in the process of determining the model. The modeling was performed on log-ratings for all tested conditions but without the level-scaled regular pass-bys as they do not represent realistic recordings. In addition to the acoustic descriptors, the type of the vehicle (variable vehicle) was encoded by 1 and 2 for the passenger car and the truck, respectively. As result, the step-wise selection resulted in the following optimal model:

$$\begin{aligned} \log_2(\text{annoyance}) = & 0.1535 \times L_{N50} - 0.06259 \times L_{Aeq} \\ & + 0.1575 \times \text{vehicle} - 9.118. \end{aligned} \quad (1)$$

In a variance analysis, the median loudness level  $L_{N50}$  explained 87% of the total variance whereas the vehicle type explained a further 4%. Consideration of the  $L_{Aeq}$  increased the explained variance by 3.6% leading to a total explained variance of 94.6%. Considering a single-descriptor model, the peak loudness level  $L_{N5}$  was able to explain 88% of the variance whereas the  $L_{Aeq}$  explained only 68%.

### 4.2.2 Urgency rating

As listeners were asked to rate the change in interior noise, analysis of the urgency ratings was done using the difference between the acoustic descriptor during a rumble strip pass-over (the first 1.2 s) and the respective quantity during regular driving. These differences are denoted using a  $\Delta$  before the quantity, e.g.  $\Delta L_{N50}$  for the difference in the median loudness level.

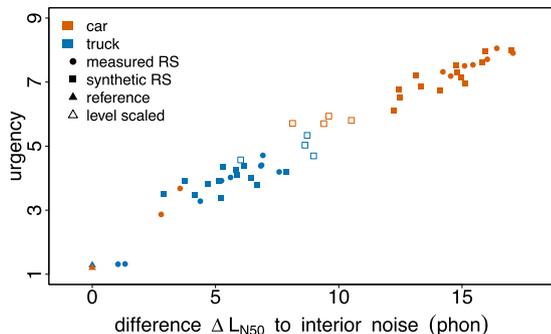
Figure 5 shows the urgency ratings as a function of  $\Delta L_{N50}$ . There is a clear relation between those two quantities and  $\Delta L_{N50}$  as a single-descriptor model was able explain 96% of the total variance in the ratings (highest value for a single variable). The step-wise model selection performed on the acoustic descriptors revealed the following optimal model:

$$\begin{aligned} \text{Urgency} = & 0.8783 \times \Delta L_{N50} - 0.3681 \times \Delta L_{N5} \\ & - 2.5252 \times \Delta S_{A5} - 0.1029 \times \Delta L_{Ceq} \\ & + 1.4675, \end{aligned} \quad (2)$$

which explained 97% of the total variance. Compared to the single-descriptor model, the optimal model cannot explain substantially more variance. For the final model, the residual standard error is 0.33, which is less than half the difference between adjacent urgency categories.

### 4.2.3 Reaction time

The probability of reacting to catch trials was essentially 0% (1 reaction out of 192 cases). On the other hand, listeners reacted almost always to all strip types with the

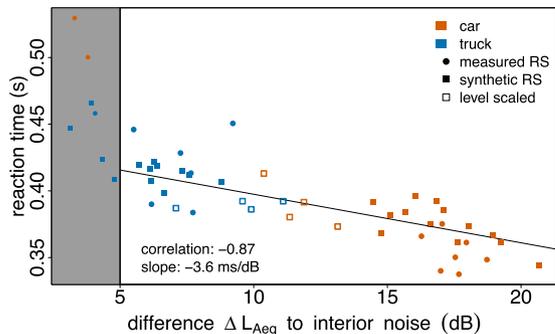


**Figure 5.** Urgency as a function of the loudness level  $\Delta L_{N50}$ . Triangles show regular pass-bys. Open symbols show the level-scaled conditions. The color denotes the vehicle type.

exception of S6r and S6 in the truck, for which the rate of reaction was below 10%. This is consistent with a rather small change in the level caused by these two variants and nearly no change in the perceived urgency for the truck. In the passenger car the probability to react to S6r and S6 was around 90–95%.

For the analysis of the reaction time, results of one listener were excluded because this listener produced considerably more misses *and* the median reaction times deviated more than threefold the standard deviation from the average of the median reaction times of all listeners. Figure 6 shows the reaction times as a function of  $L_{Aeq}$  related to the reference  $L_{Aeq}$ . The reaction times show a noisy variation in the range of milliseconds across the levels. This may originate from the uncertainty of calculating the reaction times for the recorded rumble strips, for which the moment of hitting the groove can only be determined based on the acoustic signal and thus no better than in the range of milliseconds. For the simulated data, however, we can exactly (on the basis of an audio sample, i.e., 20  $\mu$ s) determine the onset of the groove. Thus, from the reaction times obtained for the simulated strip types, a linear model of the reaction time as an effect of  $L_{Aeq}$  was derived. Only levels with a difference larger than 5 dB were included for which reactions were highly reliable. The model showed that the reaction time decreases by 3.6 ms per 1 dB of the level increase. The 95% confidence interval for the slope was  $\pm 1.1$  ms per dB. For the level changes smaller than 5 dB, the reaction times increased more strongly, being less reliable for modeling. Due to the low elicited acoustic response, this may in part also be a consequence of an increased uncertainty in detecting the exact onset.

Note that our measurements in the driving condition were done without any additional sources such as a running radio. Assuming a typical listening situation with A-levels between 70 and 80 dB, an impact of the strip in the range of 5 dB level increase without music translates to change in level of around 2 dB with the music switched on. Note that this is a purely energetic argument and that spectral masking effects may additionally reduce the impact of such rumble strips. This further underlines that a purely acoustical effect with an SPL increase in the range of 5 dB is of borderline usability in practical situations.



**Figure 6.** Reaction time as a function of the change in  $L_{Aeq}$ . The symbols show the actual reaction times as averages across all listeners. Open symbols show the level-scaled conditions. The color denotes the vehicle type. The line shows the regression between the predicted reaction times and  $L_{Aeq}$  (see text).

## 5 Acoustics: second measurement

In a second measurement, strip candidates requiring further investigations were implemented on a second test track (Fig. 7) and their acoustic emissions were recorded. The strip type C6r was implemented to see whether the tonality reduction found for pseudo-random groove distances of 300 mm also holds for 600 mm and also to see how an increased spacing and a decreased groove depth influence the SPL changes. The strip type S3 was produced to see whether the flat milling drum produces a different behavior for this strip. R4 and R6 were aimed to obtain more information on rounded variants. The sinusoidal strips S6.2, S6.5, and S6.7 were measured to provide more data for vibration measurements for strips that do not produce large acoustic responses. It is important to note that the nominal depth values of the shape wheels used for milling were actually 4, 7, and 10 mm where the latter was thought to give more insights on the effects of an increased depth on the acoustic properties. Unfortunately all three rumble strips turned out to be off by approximately 2–3 mm due to difficulties in controlling the depth.

We also were able to remeasure the interior noise for the strips from the first measurement, allowing us to check the measure-remeasure consistency.

### 5.1 Methods

#### 5.1.1 Tracks

The second test track was located on the hard shoulder, along the same highway loop but in the opposite direction of the first track. This allowed us to measure both test tracks in the same amount of time. The second test track comprised seven additional rumble strips: three strips (R4, R6, and S3) from the simulations, and four completely new strips (C6r, S6.2, S6.5, and S6.7), all shown in Figure 7. Four rumble strips were placed 30 cm from the outer road marking. Three further rumble strips were placed in parallel. Note that C6r was generated by multiplying the spacing of C3r by two and producing only half of the grooves. The width of the strips was increased to 35 cm to fully

label	R4	R6	C6r	S3	S6.2'	S6.5	S6.7
profile							
shape	rounded	rounded	cylinder	sine	sine	sine	sine
pattern							
picture							
drum	flat	flat	flat	flat	flat	flat	flat
W / mm	350	350	350	350	350	350	350
L / mm	150	200	120	-	-	-	-
S / mm	400	600	600*	300	600	600	600
D / mm	7	7	7	0-7 (10)	0-2 (4)	0-5 (7)	0-7 (10)

**Figure 7.** Additionally implemented strip types on a second track for acoustic validation.  $W$  denotes the width,  $L$  the length and  $D$  the depth of the grooves.  $S$  denotes the spacing. The strip marked with \* was only measured in the inside of the vehicle. + denotes the average spacing for irregularly spaced grooves. Depth values in parenthesis were the targeted values, values outside of the parenthesis provide the true values.

accommodate the truck front tire. The rumble strip segments were again 50 m in length. The spacing between the rumble strip sections was 30 m for the first rumble strip and slightly more than 100 m for the remaining three cross-sections. The shoulder was slightly lower ranging between 0.5 and 1 m. All other details were as described in Section 2.1.

### 5.1.2 Recordings

Generally, the recording methods followed those from Section 3. However, because of the limited availability, the conditions were restricted. The full set of positions as used for the first track was measured for R6 and R4. For S6.5, S6.7, S3, and C6r only outdoor recordings at 7.5 m were performed in parallel to the interior recordings. Thus, for these four sets outdoor recordings using microphone kit UA1404 (connected to the amplifier 2690-0S2, both Brüel & Kjaer) were performed. The signals were recorded with the multitrack recorder Edirol R4-Pro (Roland). All signals were calibrated with on-site calibration measurements (Brüel & Kjaer 4231 calibrator). Because of the different measurement setup, the light barrier could not be used. Thus, vehicle speeds were determined from either the indoor GPS measurements or the parallel outdoor measurements of R6 and R4. The time of pass-by was determined manually. The strip type S6.2 was recorded in the interior only because, based on the results of S6c and S6, no significant outdoor noise was expected for such shallow grooves.

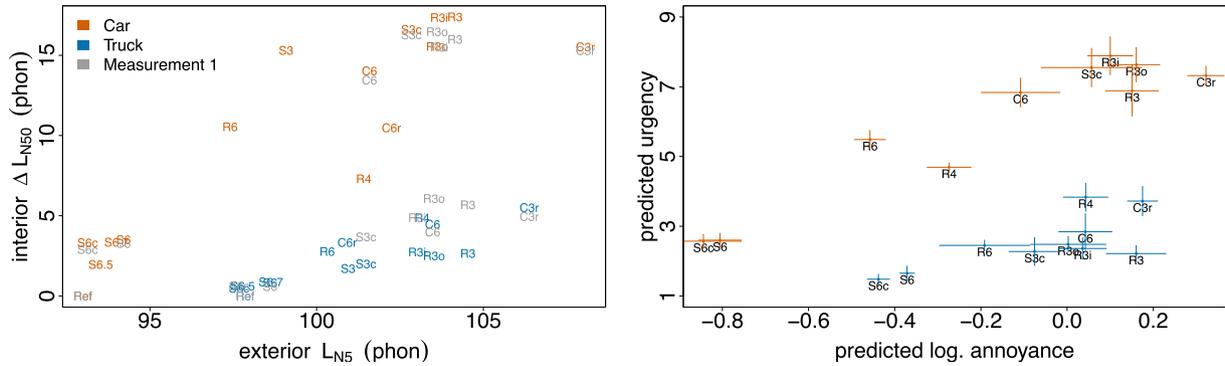
Measurement data were acquired for each vehicle and campaign in a single day, however, the two vehicle types

were measured on two separate days. For the passenger cars, for the exterior at least 16 and up to 22 (median 17) pass-overs were measured (14–15 with a median of 15 for the interior case). For the truck 15–18 (median 17) exterior recordings took place (15–17, median 16 for the interior case). The only exception were the outdoor recordings for S6.5, for which only 7 pass-overs for the car and 10 for the truck were recorded due to problems with the recording equipment.

## 5.2 Results

For the analysis, we used the loudness  $L_{N5}$  and  $L_{N50}$  as the main descriptor for the exterior and interior, respectively, because these showed significant impact explaining the perceived annoyance and urgency (88 and 96% of explained variance, respectively). The left panel in Figure 8 shows the change in median loudness  $L_{N50}$  obtained from the interior and the peak loudness  $L_{N5}$  exterior recordings. The data show interior loudness as a function of exterior loudness in order to visualize the most silent strips in the exterior with the most impact in the interior. The reference values, obtained from pass-overs without any rumbling strip, are denoted by the term “Ref”. The color encodes the two vehicle types. As in Section 3, only data for a coverage grade of 3 or better are shown.

In order to visualize the consistency between the two measurements, grey symbols replot the interior change in loudness obtained from the first measurement, see Section 3. For the passenger car, the remeasurement difference was below 1 phon, showing a good reliability when comparing the two measurements. For the truck, however, the differences



**Figure 8.** Comparison of interior and exterior acoustics. The left panel shows the change of median loudness  $\Delta L_{N50}$  for the interior vs. the peak loudness outdoors ( $L_{N5}$ , 7.5 m). Orange denotes the passenger car, blue the truck and grey shows data from the first interior measurement. The right panel shows the predicted urgency as a function of predicted annoyance (using data from 25 m). Error bars show 95% confidence intervals.

were larger, with the first measurement being louder. One explanation for the generally louder rumble strip recordings during the first recording session might be that the front tire had to be replaced between the two measurements, indicating a potential effect of the tire on the rumble strip effects, worth of further investigations in the future. Importantly, the background noise was comparable between both recordings.

The data in Figure 8 clearly show the link between interior and exterior loudness: the louder a strip in the drivers cabin, the louder the strip for the residents. Further, for the same exterior loudness, the interior effect was less pronounced in the truck than in the passenger cars. Thus, we need to distinguish between an optimal strip type for the passenger car and the truck. For the passenger car, a good rumble strips seems to be S3 (sinusoidal shape with 300 mm groove spacing and 7 mm depth) or R6 (rounded shape with 600 mm spacing and 7 mm depth) as they both produced loud interior sounds without changing the exterior much. For the truck, however, these two strip types produced a tiny change in the interior (less than 2 phon) which might be not noticeable, especially in loud interior conditions with e.g. the radio being switched on. For trucks, R4 (rounded shape with 400 mm groove spacing) seems to be a better choice. R4 is particularly interesting since in the truck it is one of the loudest strips being louder than comparable other spacings (R3, R6) but in the car it is significantly less loud than R3 and R6. The reason for this effect is, however, unclear.

Concerning the perception, the right panel of Figure 8 shows the predictions from the optimal models of urgency and annoyance using the measurements of all outdoor recordings and the interior recordings of the second measurement. As expected, the larger the annoyance, the higher the urgency caused by a strip type. The correlation coefficients between the annoyance and urgency were 0.95 for the passenger car and 0.72 for the truck. Note that S6.5, S6.7, S3, and C6r are not shown because the prediction model is based on the immission measurements at 25 m which were not performed for these rumble strips.

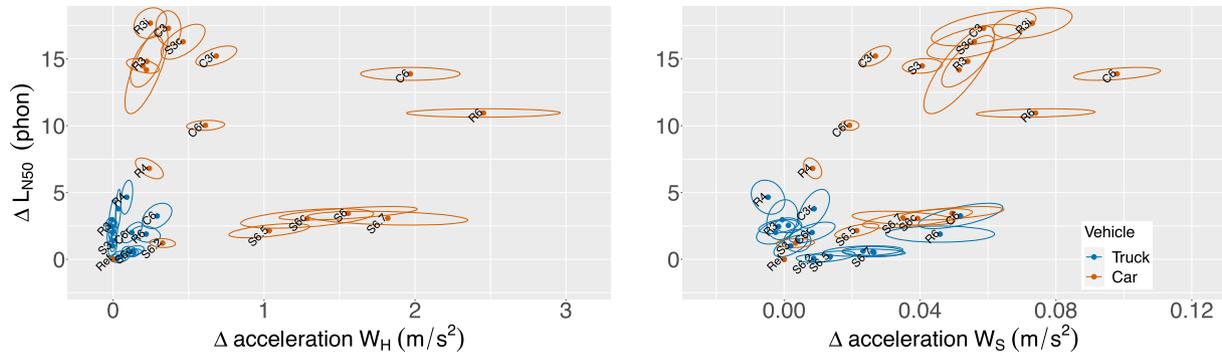
Loudness was the dominant factor for urgency and annoyance. As for the tonality, the reduction of pseudo-random spacings for 300 mm was observed for 600 mm in the interior of the car to a lower degree, i.e. the values of C6r were lower compared to C6 and R6. In the pass-by recordings the tonality values were very similar. Concerning the dependency of outdoor roughness on groove spacing and profile the additional measurements are still not entirely conclusive. C6 produces higher roughness values than R6 indicating an effect of the groove. R4, on the other side produces higher values than R6 and R3 almost in the range of C6. From the profiles it seems that R4 is less rounded than R3 and in particular R6 which may be the reason for this effect.

## 6 Vibrations

In addition to the acoustics, we analyzed the impact of the various strip types on the vibrations, looking for strip types producing large levels of vibrations, thus, well-raising attention. The measurements were performed within the acoustical measurement described in Sections 3 and 5. A tri-axial TEDS acceleration sensor (66A11 Meggitt) was mounted at the top of the steering wheel and a tri-axial whole-body seat-pad accelerometer (SV 38, Svantek, conforming with ISO 8041) was placed on the driver's seat. Both sensors were connected to the DIC-24 multi-channel system. Only data from the second measurement are reported as there all strips were recorded in a single setting and are thus directly comparable. Note that our measurements were done for two vehicles only and their parameters such as the axle distance may interact with our results.

### 6.1 Steering wheel

The vibration data from the steering wheel sensor were weighted according to ISO 5349-1. The data were high-pass filtered at 25 Hz excluding all the effects not related to the rumble strips, which were above their lowest possible frequency of around 37 Hz (for a truck at 80 km/h and



**Figure 9.** Median loudness level vs weighted vibration levels. Left panel: Vibrations at the steering wheel. Right panel: Vibrations at the seat. The circles show averages and the ellipses the 95% confidence intervals. Data are from the second measurement.

600 mm groove spacing). For the passenger car, the filtering did not affect the results by much. For the truck, the filtering reduced the weighted acceleration levels by about  $0.1\text{--}0.2\text{ ms}^{-2}$  indicating that the truck vibrations included a significant level of vibrations at frequencies below 25 Hz.

The left panel of Figure 9 shows the averages and 95% confidence ellipses for the change in interior loudness vs. the weighted acceleration obtained from the second measurement. The bivariate confidence ellipses were derived by *confidenceEllipse* from the package *car* [32] for R [21]. All strips with the groove spacing of 600 mm showed high vibration levels. For the car, the sinusoidal shapes were quite silent but still produced strong vibrations, supporting our hypothesis on these strip types. Only the extremely shallow S6.2 exhibited significantly lower vibration levels. For the truck, all changes in vibration levels were quite low, suggesting a strong damping of vibrations at the steering wheel and indicating that vibrations as an effect of running over a rumble strip are not noticeable to a driver in a truck.

## 6.2 Seat

For the seat vibrations, the ISO 2631-1 weighting was applied and the largest level along the three axes was analyzed. The vibration levels for the reference conditions, i.e., without any rumble strip, were in the range of those with the strips, caused by high levels in very low frequencies as an effect of the pavement itself. Thus, similarly to the steering wheel vibrations, high-pass filtering was applied to the seat vibrations. The filtering had a profound effect on the vibrations, reducing them from the range between  $0.1$  and  $0.4\text{ ms}^{-2}$  to the range of up to  $0.1\text{ ms}^{-2}$ .

The right panel of Figure 9 shows the change in loudness as a function of the seat vibrations. Compared to the vibrations of the steering wheel, the effects of the strip types were smaller and not as clear. For the passenger car, the louder strip types yielded more vibrations, with the 600 mm groove spacing sinusoidal strip types being in the lower range, similarly to the steering wheel vibrations. However, even for these strip types, their vertical components were above the perceptual thresholds [33]. For the truck, the vibration levels were mostly below the threshold in the relevant frequency range.

## 7 Conclusions

In this study the acoustic and vibration effects of rumble strips were investigated inside and outside a vehicle. The acoustic quantities were expressed in perceptual measures. The aim was to create a link between the acoustic and perceptual effects of rumbling in order to find ways to balance the noise burden for the residents and raising attention in the vehicle. In total, 16 different strip designs were tested on two vehicle types, a passenger car and a truck.

The combined analysis of acoustics and perception showed that loudness can well explain the perceived annoyance and urgency, with the explained variance of over 90%. Reactions to the rumbling were unreliable for strip effects below 3 dB. Level differences above 5 dB were able to well predict the reaction time. Our findings show that in the passenger car, an appropriate rumble strip can reduce the reaction time from 400 ms to 350 ms. In the truck, however, the reaction times were larger, the perceived urgency lower, and the steering wheel and seat vibrations lower, indicating that warning the truck driver by running over a rumble strip is extremely difficult if the effect on the surrounding areas is taken into account.

Our results further confirm that the exterior and interior noises are correlated. We found, however, that this correlation depends on the vehicle type, making it difficult to provide an optimal strip type for all vehicle types. Whereas for the passenger car, large loudness changes and thus salient warning signals can be achieved by a large range of strip types, for the truck, only a few distinct strip types reach level changes causing a reliable reaction.

Finally, we acquired profound knowledge about the effect of various strip types. The asymmetrical grooves R3i and R3o evoked essentially similar effects. When combined with driver reports of a slight pulling on the steering wheel, asymmetrical grooves seem to provide no advantage. The 600 mm sinusoidal strips (S6.X, S6.c) showed little acoustic response and high vibration levels in specific conditions only. The pseudo-random strip (C3r) showed reduced tonality and a large interior loudness, however, at the price of loud emissions in the exterior. By increasing the spacing and reducing the depth (C6r), the interior and exterior loudness decreased. The loud exterior can be attributed to sharp groove edges which can be potentially modified by

an edge rounding to reduce the noise. This is an important issue because people alongside roads may complain about the sudden noise when cars drive over these kind of strips. Still, the potential advantage of a reduced tonality of the pseudo-random strips comes at the price of a longsome milling process. For the passenger car, the rounded shape with 600 mm groove spacing (R6) seems to be optimal: it produced loud interior sounds and strong vibrations without changing the exterior much. Unfortunately, in the truck, R6 produced a tiny change in the interior only (less than 2 phon) which might be not noticeable, especially in loud interior conditions with e.g. the radio being switched on. In contrast, for the truck, the rounded shape with 400 mm groove spacing (R4) seems to be a better choice.

Note that our measurements were done for a single vehicle in each of the two vehicle types and the specific choice of vehicle and related parameters such as the axle distance may interact with our results, an issue to be considered when generalizing our conclusions to other vehicles even within the same vehicle type. A systematic investigation considering various vehicle parameters would be interesting for the future.

Further, parts of our investigations were based on synthesized signals. Our synthesis method seems to be of sufficient quality within our study and allowed us to form a better understanding of the relation between perception and acoustics of rumble strips. However, we encountered issues beyond the frequency range of the harmonic excitation, which may reduce the generalizability of our method to other vehicle types. In order to generalize our synthesis method to a larger set of vehicle and strip types, further validations are required in the future.

## Acknowledgments

We want to thank the Austrian Research Promotion Agency, the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation, and Technology, as well as the ASFINAG for funding this project (FFG, project 850538). We also want to thank our project partner ABF Strassensanierungs GmbH for their valuable input and experience and for producing all the rumble strips according to our specifications.

## References

1. D.J. Miles, D.M. Finley: Factors that influence the effectiveness of rumble strip design. *Transportation Research Record: Journal of the Transportation Research Board* 2030 (2007) 1–9.
2. J. Kragh, B. Andersen, T.N. Sigurd: Low noise rumble strips on roads – A pilot study. In: *Proceedings of the Internoise 2007* (2007) 755–764.
3. O. Houix, S. Bonnot, F. Vienne, B. Vericel, L. Pardo, N. Misdariis, P. Susini: Perceptual influence of the vibratory component on the audio component of alarms produced by rumble strips, by measuring reaction times, in *Proceedings of the Acoustics 2012 Nantes Conference*, 2012, pp. 1083–1088.
4. E. Donnell, H. Sommer, P.M. Garvey, S.C. Himes, D.J. Torbic: Statistical model of in-vehicle sound generated from highway rumble strips. *International Journal of Vehicle Noise and Vibration* 5, 4 (2009) 308–328.
5. T.D. Blumenthal, W.K. Berg: Stimulus rise time, intensity, and bandwidth effects on acoustic startle amplitude and probability. *Psychophysiology* 23, 6 (1986) 635–641.
6. Delaware Department of Transportation: SR 24 Longitudinal edge line rumble strip noise study, 2012. [https://deldot.gov/Programs/DSHSP/pdfs/RumbleStrip\\_compressed.pdf](https://deldot.gov/Programs/DSHSP/pdfs/RumbleStrip_compressed.pdf). Last visited 3rd Dec. 2021
7. G. Watts, R. Stait, N. Godfrey, L. Chinn, R. Layfield: Development of a novel traffic calming surface “rippleprint” prepared for charging and local transport division, Transport Research Laboratory (TRL), Crowthorne, Berkshire, UK, 2002.
8. P. Donovan, B. Rymer: Design and evaluation of quieter highway rumble strips 1 (2013) 605–621.
9. International Organization for Standardization, ISO 1996–2:2007: Acoustics – description, measurement and assessment of environmental noise – part 2: determination of environmental noise levels, 2007.
10. Deutsches Institut für Normung, DIN 45681: Acoustics – Determination of tonal components of noise and determination of a tone adjustment for the assessment of noise immissions, 2005.
11. A. Hegewald, A. Vesper, M. Irzik, R. Krautscheid, K. Sander, A. Lorenzen, U. Löffler, O. Ripke, F. Bommert, Sicherheitswirkung, Dauerhaftigkeit und Lärmemission von eingefrästen Rüttelstreifen [Safety effect, durability and noise emission of milled rumble strips], Fachverlag NW in der Carl Ed. Schünemann KG, Bremen, Germany, 2018.
12. P. Donovan, The influence of tires on rumble strip noise and vibration, in *INTER-NOISE 2019 Madrid – 48th International Congress and Exhibition on Noise Control Engineering*, 2019, pp. 984–995.
13. C.H. Kasess, T. Maly, P. Majdak, H. Waubke: The relation between psychoacoustical factors and annoyance under different noise reduction conditions for railway noise. *The Journal of the Acoustical Society of America* 141, 5 (2017) 3151–3163.
14. M.E. Nilsson, M. Andéhn, P. Lešna: Evaluating roadside noise barriers using an annoyance-reduction criterion. *The Journal of the Acoustical Society of America* 124, 6 (2008) 3561–3567.
15. R.B. Raggam, M. Cik, R.R. Höldrich, K. Fallast, E. Gallasch, M. Fend, A. Lackner, E. Marth: Personal noise ranking of road traffic: subjective estimation versus physiological parameters under laboratory conditions. *International Journal of Hygiene and Environmental Health* 210, 2 (2007) 97–105.
16. M. Müller, A. Telle, A. Fiebig, Psychoakustische Wirkung von Fahrbahnmarkierungen [Psychoacoustic effect of road markings], Fachverlag NW in der Carl Ed. Schünemann KG, Bremen, Germany, 2015.
17. J. Edworthy, S. Loxley, I. Dennis: Improving auditory warning design: Relationship between warning sound parameters and perceived urgency. *The Journal of the Human Factors* 33, 2 (1991) 205–231.
18. C. Suied, P. Susini, S. McAdams: Evaluating warning sound urgency with reaction times. *Journal of Experimental Psychology: Applied* 14, 3 (2008) 201–212.
19. Austrian Standards, ÖNORM B 3584–1:2018: Asphaltmischgut–Mischgutanforderungen–Teil 1: Splittmastixasphalt–Empirischer Ansatz–Regeln zur Umsetzung der ÖNORM EN 13108–5 [Bituminous mixtures–Material specifications–Part 1: Stone mastic asphalt–Empirical approach–Rules for the implementation of ÖNORM EN 13108–5], 2018.

20. A. Weninger-Vycudil, P. Simanek, T. Rohringer, J. Haberl: Handbuch Pavement Management in Österreich 2009[Handbook Pavement Management in Austria 2009]. Schriftenreihe Stra. betavenforschung des Bundesministeriums für Verkehr, Innovation und Technologie 584 (2009) 120.
21. R Core Team: Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. 2019.
22. Deutsches Institut für Normung: DIN 45631/A1: Calculation of loudness level and loudness from the sound spectrum – Zwicker method – Amendment 1: Calculation of the loudness of time-variant sound, 2010.
23. Deutsches Institut für Normung: DIN 45692: Measurement technique for the simulation of the auditory sensation of sharpness, 2009.
24. W. Aures: Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse [A procedure for calculating the consonance of any sound, a contribution to auditory sound analysis]. Ph.D. Thesis, Technical University of Munich 1984.
25. R. Sottek, Modelle zur Signalverarbeitung im menschlichen Gehör [Signal processing models for the human auditory system]. Ph.D. Thesis, RWTH Aachen, 1993.
26. E. Terhardt, G. Stoll, M. Seewann: Algorithm for extraction of pitch and pitch salience from complex tonal signals. *Journal of the Acoustical Society of America* 71, 3 (1982) 679–688.
27. J. Vos: Annoyance caused by the sounds of a magnetic levitation train. *The Journal of the Acoustical Society of America* 115, 4 (2004) 1597–1608.
28. C.H. Kasess, A. Noll, P. Majdak, H. Waubke: Effect of train type on annoyance and acoustic features of the rolling noise. *The Journal of the Acoustical Society of America* 134, 2 (2013) 1071–1081.
29. H. Lane, A. Catania, S. Stevens: Voice level: Autophonic scale, perceived loudness, and effects of sidetone. *Journal of the Acoustical Society of America* 33, 2 (1961) 160–167.
30. G.E. Schwarz: Estimating the dimension of a model. *Annals of Statistics* 6, 2 (1978) 461–464.
31. W.N. Venables, B.D. Ripley: *Modern applied statistics with S* (4th ed.), Springer, New York, 2002.
32. J. Fox, S. Weisberg: *An R companion to applied regression* (3rd ed.), Sage, Thousand Oaks, CA, 2019.
33. I. Baumann, M.A. Bellmann, V. Mellert, R. Weber: Wahrnehmungs- und Unterschiedsschwellen von Vibrationen auf einem Kraftfahrzeugsitz [Perception- and difference thresholds of vibrations on a driver's seat]. *Fortschritte der Akustik – DAGA 1* (2001) 292–293.
34. D. Duhamel: Efficient calculation of the three-dimensional sound pressure field around a noise barrier. *Journal of Sound and Vibration* 197, 5 (1996) 547–571.
35. C.H. Kasess, W. Kreuzer, H. Waubke: Deriving correction functions to model the efficiency of noise barriers with complex shapes using boundary element simulations. *Applied Acoustics* 102 (2016) 88–99.
36. R.A. Broadbent, D.J. Thompson, C.J.C. Jones: The acoustic properties of railway ballast. *Euronoise* (2009) 3307–3316.
37. H. Waubke, C.H. Kasess: Simulation of the noise radiation and shielding with boundary element method (BEM), in *Proceedings of the ICSV23*, 2016.
38. International Organization for Standardization: ISO 3095: 2013 Acoustics – Railway applications – Measurement of noise emitted by railbound vehicles, 2013.
39. C.H. Kasess, H. Waubke, M. Conter, C. Kirisits, R. Wehr, H. Ziegelwanger: The effect of railway platforms and platform canopies on sound propagation. *Applied Acoustics* 151 (2019) 137–152.
40. J. Defrance, P. Jean: Integration of the efficiency of noise barrier caps in a 3D ray tracing method. Case of a T-shaped diffracting device. *Applied Acoustics* 64, 8 (2003) 765–780.
41. C.H. Kasess, H. Waubke: Moving sources and the 2.5D Helmholtz boundary element method, in *Proceedings International Congress on Acoustics*, Aachen, 2019.
42. D. Marelli, P. Balazs: On pole-zero model estimation methods minimizing a logarithmic criterion for speech analysis. *IEEE Transactions on Audio, Speech, and Language Processing* 18, 2 (2010) 237–248.

## Appendix

An important part of this work required extending the set of rumble strips data for those strip types that we were not be able to manufacture on the test track. For this, measurement data were simulated and these simulations were used in the main article, see [Section 4.1.2](#). This section provides a detailed description of the simulation approach used to generate synthetic rumble strip pass-overs.

### A.1 Outdoor simulations

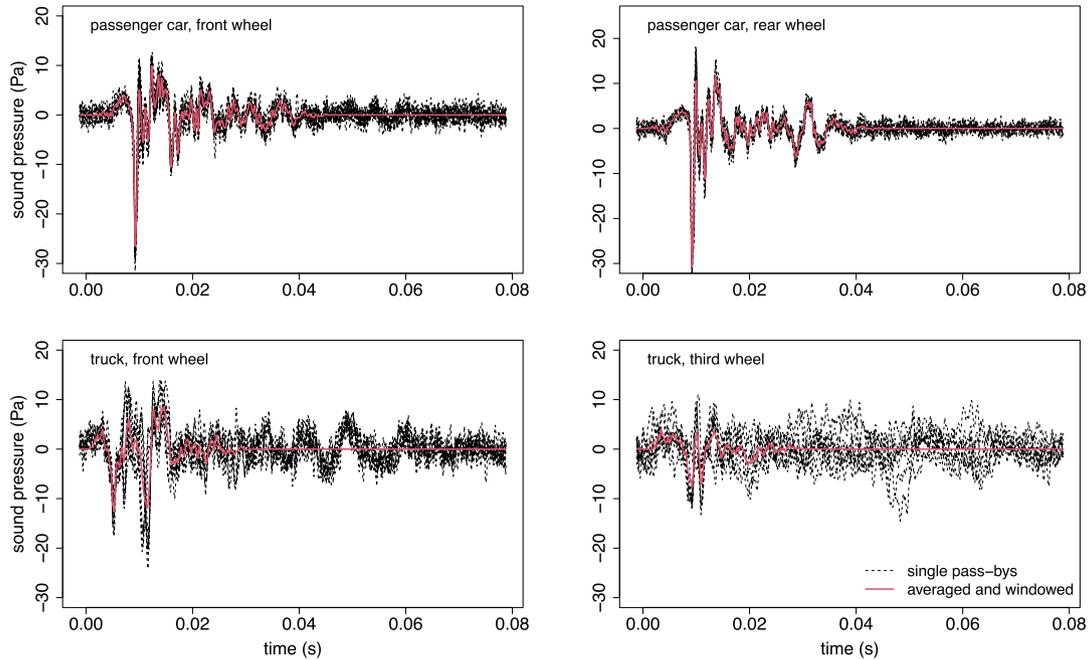
The simulated outdoor stimuli were used to extend the set of stimuli used in the annoyance rating experiment (described in [Sect. 4.1.3](#)). Our simulation relies on the basic building blocks similar to those from [16]. First, a reference signal for a single groove is generated from measurements. Then, using the vehicle motion and groove positioning a source signal for the pass-over is produced which may also be modified depending on the groove shape. Last, this signal is propagated into the surroundings based on the desired measurement position and the vehicle motion resulting in the signal to be presented.

As a basis for the synthetic outdoor rumble strip noise the single-groove measurements were used. For the single groove measurement a microphone was placed 1 m away from the edge and 0.315 m above the road surface.

Along the first two strips (C3 and C6), a single groove was placed (depth 12 mm) that was used to measure the impulse response. Two further single grooves were placed with an inward and outward directed slant (see conical design). As the conical strips R3i and R3o produced similar acoustic emissions as the regular strip R3, the latter two were, however, not used to produce synthetic stimuli but only to evaluate the propagation (see [Sect. A.1.2](#)).

#### A.1.1 Car and truck noises

For the car, [Figure A.1](#) shows the measurements of the front wheel (upper left panel) and the rear wheel (upper right panel) for a tire coverage grade of 2 or less (9 out of 11). As a reminder, 0 implies a perfect hit and  $\pm 5$  a complete miss. The different pass-overs were manually synchronized based on the highest peak response. Clearly, these 9 pass-overs produced highly reproducible excitation signals. This is also due to a consistently high tire coverage. Only 2 out of 11 pass-overs were slightly off and there were



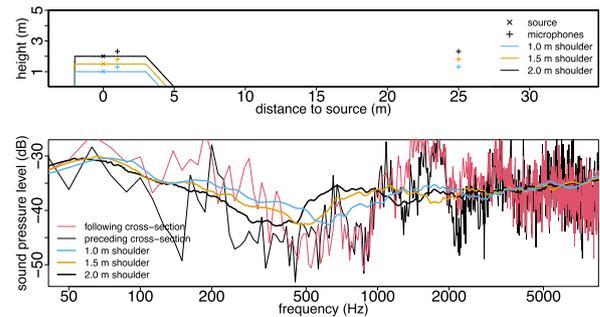
**Figure A.1.** Single-groove measurements for the passenger car and truck. Shown are single pass-overs (black dashed lines) for a good coverage (lower than 2) and the averaged and windowed responses (red line). The upper panels show the front (left) and the rear tire (right) of the passenger car, the lower panels shown the front (left) and the third tire (right) of the truck.

essentially no clear misses. To reduce the effect of background noise on the synthesis an average response was generated (red lines in Fig. A.2), excluding the two aforementioned pass-overs. Furthermore, a window function was applied to the average response to reduce the effect of background noise at the tails of the response.

For the truck, Figure A.1 shows the measurements from the front wheel (lower left panel) and the rear, i.e. the third axle (lower right panels) for good front tire coverage (2 or less). The middle tire is very similar to the rear tire and is thus not shown. Clearly, in particular the rear wheels exhibit relatively low signals and poor reproducibility. Here, it is important to keep in mind that only the front tire coverage was monitored, thus the higher spread for the rear axles may be due to varying degrees of coverage. Furthermore, the truck has wider tires in general and the second axle is equipped with twin tires. Due to the high background noise of the engine, the averaging of the responses (red lines in lower panels of Fig. A.1) was even more important than for the car. To reduce the amount of smearing for the average response only pass-overs with coverages equal or better than 2 were used. Again a window was applied to all axles.

### A.1.2 Noise propagation

For the synthesis the signal was needed at a distance of 25 m from and a height of 1.2 m above the rumble strip. To model the noise propagation, the 2.5D boundary element method (BEM) was used [34, 35] which assumes a constant cross-section. As the height of the shoulder with respect to the surroundings varied along the test track, three different



**Figure A.2.** Simulated cross-sections. The upper panel shows the different cross-sections calculated using the BEM, distinguished by color. The main difference is the height of the shoulder above the surroundings. The x at 0 m indicates the position of the tire, the crosses indicate the microphone positions close to the groove and at immission distance. The lower panel shows the comparison of the measured and the static simulated transfer function. The transfer function is defined here as the spectral difference of the single-groove measurements at 1 m and 25 m, the latter in the two adjacent rumble strip cross-sections.

cross-sections were calculated (upper panel in Fig. A.2) with shoulder heights of 1 m, 1.5 m, and 2 m above the ground.

The ground was modeled using a mildly absorbing impedance model [36]. This three-layer model was derived for ballasted railway tracks, however, due to the multiple layers it avoids problems with interferences as in hard-backed single-layer models (cf. [37]). Furthermore, simulations using a standard ground model lead to narrow-band effects and numerical problems. The lower panel in Figure A.2

shows the different transfer functions from 1 m to 25 m distance for the BEM simulations (thick lines) and from the measurements for the 25 m position. Since the 25 m were not measured in the single-groove cross-section, to obtain an estimate of the propagation, the single-groove events measured in the preceding and the following cross-section of the single-groove pass-overs were used. The three single grooves which were placed between first and second, second and third, and third and fourth rumble strip were used. Thus, e.g. while measuring C6, the good pass-overs of the single groove preceding and following C6 could be used to obtain an estimate of the transfer function. For this, only the passenger car data were used as the engine noise was lower. The signals of good hits were manually aligned in time (front and back separately), averaged and an FFT was applied (0.08 s window length). The resulting spectra were compared to the spectra of the 1 m measurements of the corresponding single grooves. The frequency shift due to the motion of the vehicle was corrected using the standard Doppler shift for the speed and angle of motion. Although there were some differences, the simulations are on average in good agreement. The data at 25 m are measured from almost 50 m distance, so the noisiness is to be expected. The most likely reasons for the difference between the preceding and following cross-section are a potential directivity of the source in driving direction as well as the steadily increasing height of the shoulder compared to the surroundings. A similar shift of the dip between 100 and 1000 Hz can be observed in the simulations when changing the shoulder height.

Air damping was included by simply using the path length from source to receiver as a basis for the amplitude attenuation according to ISO 9613-1 [38–40].

### A.1.3 Moving sources

To consider the moving source a simplified approach was chosen where the response of the tire was assumed to be local at the according position by multiplying the source spectrum with the sum of the transfer functions at the points of the grooves (and applying the proper phase shifts to accommodate the different time shifts). This is of course not entirely correct as the tire is moving during the excitation which also shifts the excitation spectrum itself in the receiver location. However, accommodating these effects is difficult as the tire itself is a rotating source and it seems unclear how to incorporate the exact motion of the excitation as different points of the tire move at different speeds relative to the stationary observer. Furthermore, the tire moves in the range of 1 m during the excitation and thus in 25 m the change of position does amount to a maximum change in angle of only about 2–3 degrees which is a relatively small change. Finally, the computational effort when taking the motion during the excitation into account is considerably higher.

The calculation of the transfer function for the pass-overs requires a resolution of 0.4 Hz in order to generate the required signal duration of 2.5 s for the 50 m long

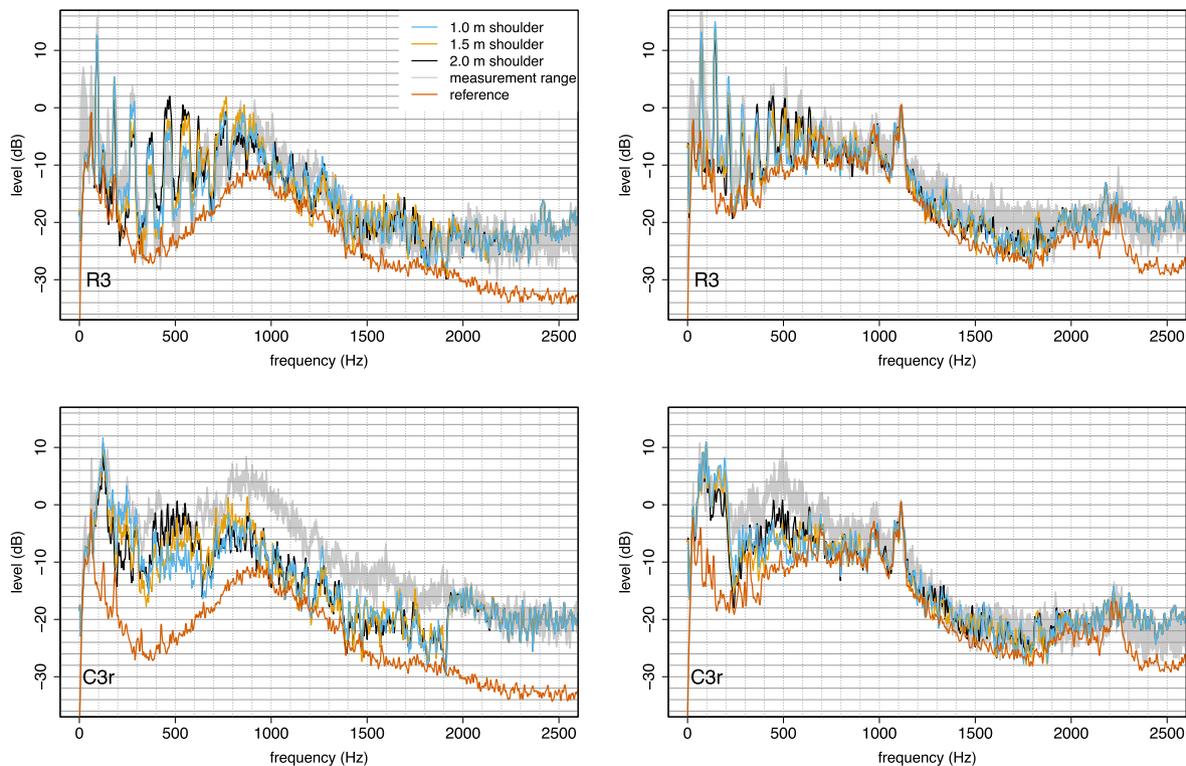
rumble strip. BEM calculations were performed up to 8 kHz which in turn implies a total of 20,000 3D frequencies to be calculated. Even using an efficient scheme such as in [35] leads to 2D BEM calculations in the range of  $10^6$  which is clearly not feasible. Thus, a spectral interpolation was performed by demodulating the transfer function using the source-receiver distance and the interpolating between the calculated frequencies (for a description see [41]).

To include binaural cues (i.e. the interaural time difference and interaural level difference) in the synthetic rumble strip noise the results for a plane wave scattering on a sphere were used. The direction of the plane wave was determined according to the position of the respective moving sound source.

### A.1.4 Groove parameters

The various depth profiles were simulated by assuming that smoother geometries lead to less high-frequency contributions. In order to encompass this effect, spectral envelopes from the measurements of the different rumble strips were extracted. The method used is similar to the method described for speech in [42]. First, the signal was cut such that roughly 10–11 m around the measurement cross-section were covered. This reduces the influence of the changes in the transfer function as well as the amount of Doppler smearing. The signal was downsampled to 4 kHz as harmonic excitation was limited to below 2 kHz. Then, a Hanning window of 150 ms (hop size 10 ms) was used to calculate the DFT. The fundamental frequency was determined, depending on the speed and the rumble strip spacing and the harmonic peaks detected. The spectral energy of the pass-by noise without a rumble strip was energetically subtracted. In case of a resulting negative energy, the energy was set to zero. After this the corrected peak amplitudes at the harmonics were determined. In case the amplitude was more than 10 dB lower than the regular pass-by amplitude, the pass-by amplitude was used. This procedure was done to avoid problems with low signal amplitudes in the rumble strip pass-over which would lead to noisy estimates of the rumble strip differences. Using the peak amplitudes, a cubic spline was fitted through the log-amplitudes of these peaks.

Zero-pole filters were determined from the log-envelopes with 30 zeros and 30 poles [42] to introduce a degree of smoothing to the envelope. To determine the difference R3 was compared with S3c and C6 with S6c, as these were the corresponding regular-sinusoidal pairs with similar groove spacing. Choosing the 5 best pass-overs with respect to tire coverage the median of the 25 time-averaged pairwise differences of the envelopes of the filters was calculated. Than another filter was derived as described above using the resulting spectral difference. For the synthesis of the 400 mm spacing, the weighted difference between the 300 mm and 600 mm difference curves was determined (2/3 and 1/3 respectively) and yet another filter derived. This was deemed necessary as the groove effect seemed to be highly dependent on the spacing.



**Figure A.3.** Comparison of measurement and synthesis for the car (left panels) and the truck (right panels) for two rumble strips. Shown are mean, minimum, and maximum of 5 pass-bys (gray shaded area) as well as the synthesis based on the three cross-sections of different shoulder height (colored lines). The upper panel shows R3, the lower shows C3r. The orange line shows a regular pass-by.

### A.1.5 High-frequency noise and level calibration

Outdoor recordings showed a strong contribution to the rumble strip noise above 2000 Hz which was not observed in the single-groove recordings. This high-frequency contribution was added to the synthetic stimuli, according to groove spacing and groove type. The magnitude of the increase was quantified using the bandpass filtered noise (2000 Hz to 8000 Hz) of the three best pass-overs compared to the regular pass-by. In the case of the 400 mm synthetic rumble strip the noise caused at 300 mm was used, the level however was adjusted to a weighted sum of 300 mm and 600 mm of the corresponding groove geometry (R3, C6 for a regular groove; S6c, S3c for a sinusoidal groove design). For the pseudo-random pattern at different average spacings the high-frequency noise of the C3r was used. For the synthetic stimuli, the rumble strip recording was highpass-filtered at 1920 Hz to avoid a dip at 2000 Hz whereas the regular pass-by was low-pass filtered at 2000 Hz.

To calibrate the level of the simulated pass-overs, the level of the synthesis and the measurement of R3 below 1000 Hz were compared as in this frequency range the rumble strip signal is dominant. The difference between these two quantities was used to adjust all synthetic rumble strip signals. The car was attenuated by about 2 dB, the truck was amplified by about 2.5 dB based on the synthesis using the cross-section with a shoulder of 1 m height.

### A.1.6 Results

The upper left panel in [Figure A.3](#) illustrates the comparison of the spectrum of the synthetic and the range of 5 measurements of R3 with good coverage for the car (gray shaded area). Different shoulder heights are color coded. The orange line illustrate a regular pass-by. The effect of the surrounding terrain on the propagation in the mid-frequency range can be clearly seen. From the comparison the 1 m high shoulder seems to most adequately model the measurements which is also roughly in agreement with the true height in that cross-section. All together the simulation is in good agreement with the measurement. For the 600 mm conventional pattern (C6, not shown here) there are some deviations in the range from 1000 to 1500 Hz which may be a consequence of the slightly less smooth and deeper groove profile compared to the R3. The larger deviations for the pseudo-random pattern (C3r) in that range support this assumption. Even though this effect lead to a discrepancy for the C3r between low and high frequency range, the synthesized C3r was still used. The sinusoidal strips are essentially in good agreement, with a slight overestimation of some of the harmonics. For the two strips of the validation track which were measured in 25 m there is some overestimation for 600 mm below 1000 Hz and above 2000 Hz but the 400 mm strip shows good agreement for the rumble strip noise, however the high-frequency component was assumed to be higher than it actually was.

For the truck (right panels in Fig. A.3) the tendencies are similar, however the agreement overall is better than for the passenger car which may partially be attributed to the higher pass-by noise of the truck which reduces the additional acoustic effects of the rumble strips. For the sinusoidal strips (S6c, S3c) and the validation strips (R6 and R4) there is some overestimation of the harmonics below 400 Hz and for 400 mm (R4) some underestimation around 500 Hz.

## A.2 Synthetic interior noise

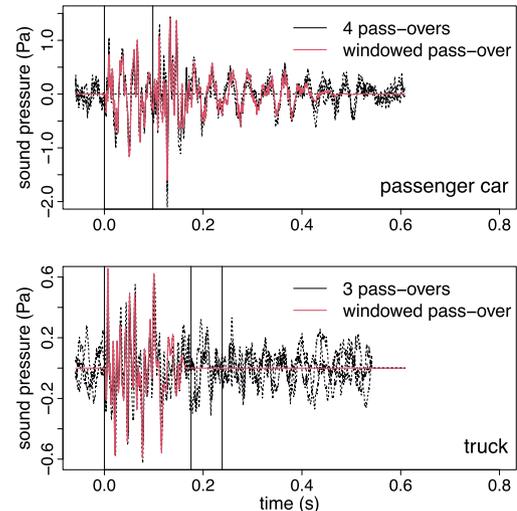
To extend the set of interior noises for the urgency rating and the reaction time test (described in Sects. 4.1.4 and 4.1.5, respectively), different synthetic noises were generated. In contrast to the outdoor noise, for the synthetic interior noise a different strategy was used for a number of reasons. For one, there was no need to calculate propagation paths, as the microphone positions are the same for the response and the listener. However, in the course of extracting the response to a single groove it turned out that the signal-to-noise ratio due to the interior noise was poor, especially in the truck where the response was extremely weak and barely audible. Thus, instead of extracting the single groove response, an inverse filtering was performed on the measurements of the pseudo-random strip. The HATS data were used to also get the spatial information encoded in the two channels.

### A.2.1 Details

Using the known groove pattern and the vehicles' speed a temporal impulse train was generated. The speed was varied between 95 km/h and 105 km/h (78–83 km/h for the truck) in steps of 0.1 km/h to achieve the best correlation with the measurement. To reduce the data volume and the calculation time, the audio signal was downsampled to 4 kHz as the harmonic excitation happened below 2 kHz. The solution for a linear system of equations relating the unknown response to the measured signal via the pulse train was calculated via the pseudoinverse (function *pinv* in Matlab) for each recording channel of the HATS separately. The pass-over with the best overall coverage (root mean square or RMS) was used.

One difficulty for the car was that the response of the front and back tire could not be separated with this approach. Since the speed of the vehicles was not altered for the synthetic noise this did not pose a problem. The upper panel of Figure A.4 shows the responses of 4 pass-overs determined using the inverse filtering approach. Clearly, there is a slowly decaying low frequency component that extends to at least 0.5 s. A window was used to smoothly fade out this component.

For the truck no noticeable response of the rear wheels was observed. Furthermore, there seems to be low frequency components, however, in contrast to the car these are out of phase and thus were assumed to be unrelated to the response of the truck to the groove. As a consequence, the response was cut appropriately to reduce the effect of regular cabine noise.



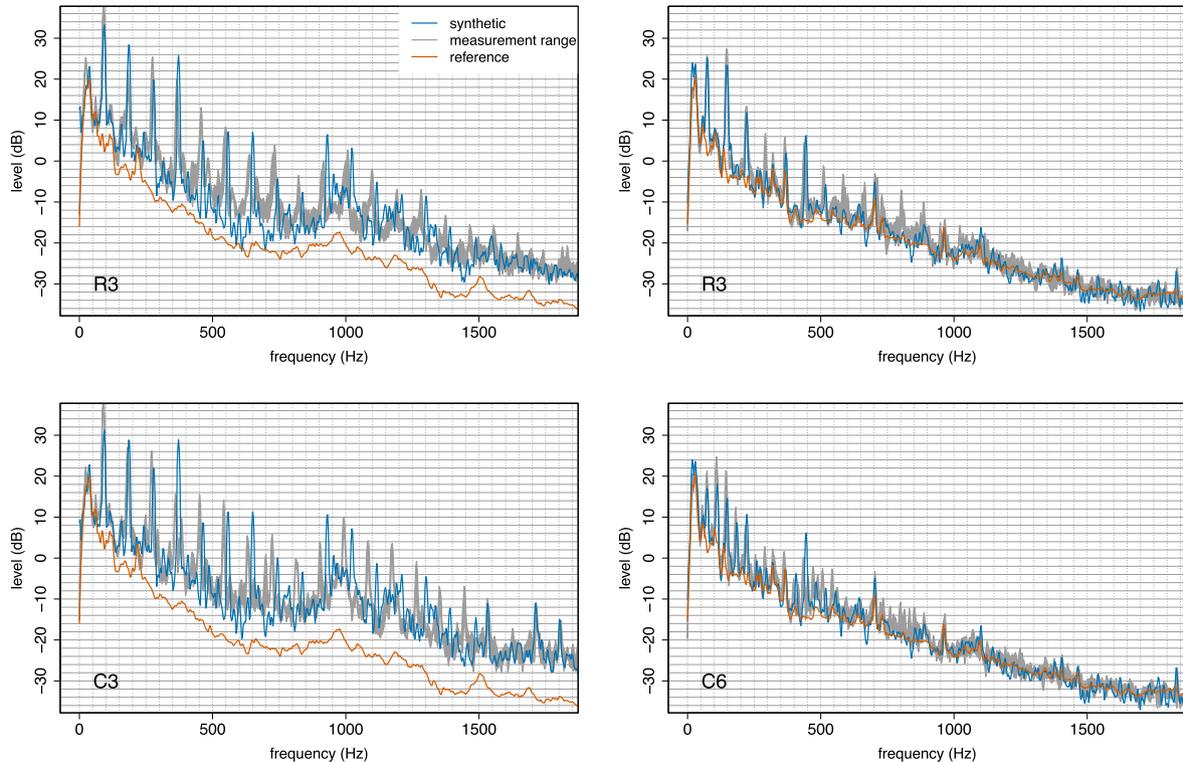
**Figure A.4.** Inverse filtering results. Shown are pass-overs with a good average coverage for the passenger car (black lines). The vertical lines mark the axles. The red line shows the windowed selected response for the passenger car. The lower panel shows the same for the truck.

Similar to the exterior noise, a pulse train based on the groove pattern and the speed was generated and convolved with the respective impulse responses of the ears of the HATS.

The rumble strip profile was taken into account in a way similar to the exterior synthetic noise, there were, however, some differences. First, the whole pass-over was used. As the rumble strip signal was stationary inside the vehicles, a longer Hanning window of 300 ms (hop size 10 ms) was used. Furthermore, the background noise was not considered as the harmonics were clearly dominant. For the truck the frequency range was limited to 1500 Hz instead of 2000 Hz as no contribution of the rumble strip was found above that frequency. For the car, the synthesized 300 mm spacing was closest to the measured R3. Based on this finding, the difference between the synthesized and measured C3r was used to model the difference between rounded and unrounded grooves for the car. This was necessary as R3 and C3 had slightly different true spacings which in the case of the interior noise resulted in non-overlapping narrow-band harmonics. For the car these were relatively widely spaced and resulted in problematic effect in the difference of the envelope. Other geometries were produced as in the exterior case based on R3. For the truck, the synthesis using the original impulse responses were more closely related to C3 which was used as a basis after correction of the difference between measured and synthesized C3r. Deriving the difference between R3 and C3 caused no problems for the truck, as the harmonics are less far apart.

Similar to the exterior noise, signals above 2000 Hz were for the car adapted as described. For the truck no such step was required as no contributions above 1500 Hz were observed.

Calibration was done as for the exterior synthetic stimuli using modeled and measured levels below 1000 Hz



**Figure A.5.** Comparison of measurement and synthesis for the interior of the car (left panels) and the truck (right panels) for two rumble strips. Shown are mean, minimum, and maximum of 5 pass-bys (black lines and gray shaded area) as well as the synthesis (blue line). The upper panels show R3, the lower panels show C3 (C6 for the truck). The orange line shows a regular pass-by.

of R3. The car was amplified by about 2.1 dB, the truck was amplified by about 1.3 dB.

### A.2.2 Results

The upper left panel in Figure A.5 illustrates the comparison of the spectrum of the synthetic and the range of 5 measurements of R3 with good coverage for the car (gray shaded area). The orange line illustrates the regular background noise. There is an overall good agreement between measurement and synthesis although some harmonics differ considerably. For C3 (lower left panel) a similar tendency was observed, with the 3rd harmonic being extremely overestimated. In the same frequency range, C6 and S3c also shows an overestimation of one harmonic. Looking at the higher harmonics, it becomes clear that

there is a mismatch in the fundamental frequency meaning that the groove spacing is slightly off. When simulating C3 with 307 mm instead of 300 mm, the agreement is much better indicating a high sensitivity with respect to the groove pattern. Still, for the synthesis the rounded values (300, 400, and 600 mm) were chosen irrespective of the true spacing.

For the truck (right panels in Fig. A.5) the tendencies are similar, however the agreement overall is better than for the passenger car. Again, some harmonics are off, but to a lesser degree than for the car.

For both vehicles the sinusoidal strip with 600 mm spacing was not considered in the synthesis, as the effect of the rumble strip was extremely low, or in the case of the truck, inaudible.

**Cite this article as:** Kasess C.H. Maly T. Majdak P. & Waubke H.. 2021. Effects of rumble strip parameters on acoustics, vibration, and perception. Acta Acustica, 5, 54.