A robust sequential hypothesis testing method for brake squeal localisation

Nilesh Madhu, Sebastian Gergen, and Rainer Martin

ARTICLES YOU MAY BE INTERESTED IN

Noise robust bird call localisation using the generalised cross-correlation with phase transform in the wavelet domain
The Journal of the Acoustical Society of America **146**, 4650 (2019); https://doi.org/10.1121/1.5138593

Time domain imaging of extended transient noise sources using phase coherence
The Journal of the Acoustical Society of America **146**, 4851 (2019); https://doi.org/10.1121/1.5138926

Introduction to the Special Issue on Acoustic Source Localization
The Journal of the Acoustical Society of America **146**, 4647 (2019); https://doi.org/10.1121/1.5140997

Estimating the spatial distribution of vocalizing animals from ambient sound spectra using widely spaced recorder arrays and inverse modelling
The Journal of the Acoustical Society of America **146**, 4699 (2019); https://doi.org/10.1121/1.5139406

Model-based Bayesian direction of arrival analysis for sound sources using a spherical microphone array
The Journal of the Acoustical Society of America **146**, 4936 (2019); https://doi.org/10.1121/1.5138126

Acoustic localization of crows in pre-roost aggregations
The Journal of the Acoustical Society of America **146**, 4664 (2019); https://doi.org/10.1121/1.5138133
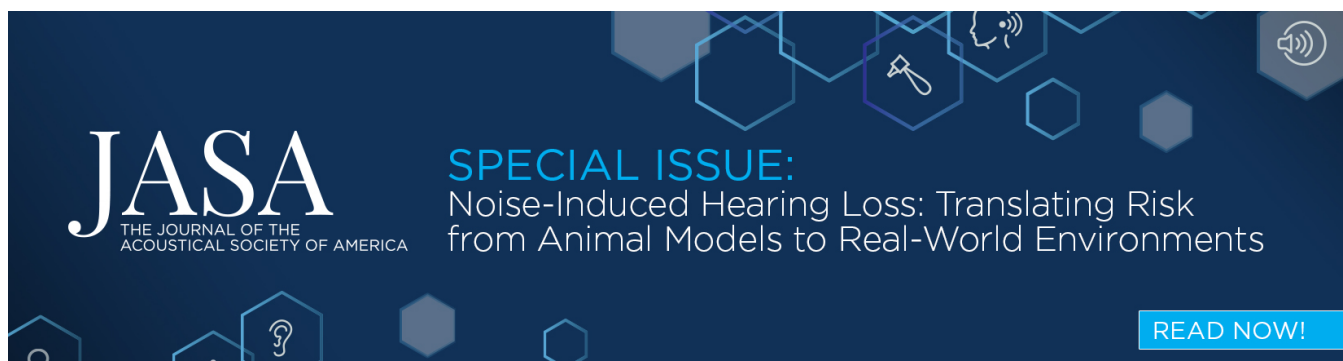
A robust sequential hypothesis testing method for brake squeal localisation[a)]

Nilesh Madhu,[1,b)] Sebastian Gergen,[2] and Rainer Martin[3]

[1]*IDLab, Department of Electronics and Information Systems, Ghent University—imec, Ghent, Belgium*
[2]*Bochum Institute of Technology gGmbH, Bochum, Germany*
[3]*Institute of Communication Acoustics, Ruhr-Universität Bochum, Germany*

This contribution deals with the *in situ* detection and localisation of brake squeal in an automobile. As brake squeal is emitted from regions known *a priori*, i.e., near the wheels, the localisation is treated as a hypothesis testing problem. Distributed microphone *arrays*, situated under the automobile, are used to capture the directional properties of the sound field generated by a squealing brake. The spatial characteristics of the sampled sound field is then used to formulate the hypothesis tests. However, in contrast to standard hypothesis testing approaches of this kind, the propagation environment is complex and time-varying. Coupled with inaccuracies in the knowledge of the sensor and source positions as well as sensor gain mismatches, modelling the sound field is difficult and standard approaches fail in this case. A previously proposed approach implicitly tried to account for such incomplete system knowledge and was based on *ad hoc* likelihood formulations. The current paper builds upon this approach and proposes a second approach, based on more solid theoretical foundations, that can systematically account for the model uncertainties. Results from tests in a real setting show that the proposed approach is more consistent than the prior state-of-the-art. In both approaches, the tasks of detection and localisation are decoupled for complexity reasons. The localisation (hypothesis testing) is subject to a prior detection of brake squeal and identification of the squeal frequencies. The approaches used for the detection and identification of squeal frequencies are also presented. The paper, further, briefly addresses some practical issues related to array design and placement. © 2019 Author(s). *All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).* https://doi.org/10.1121/1.5138608

I. INTRODUCTION

The optimisation of the acoustic design of cars over the past years has led to significant reduction in the engine-, wind- and tyre-noise, chassis vibration noise, etc. Due to this lowering of the overall sound level, a variety of low-level contributors to the acoustic footprint are now coming into the focus of the automobile manufacturers. One such contributor is brake squeal—the high-pitched noise sometimes emitted when brakes are applied. This is known to be very annoying and is a major concern for the manufacturers. The *in situ* detection and localisation of brake squeal is, however, a challenging problem since the scenario in which brake squeal occurs is a dynamic one: the background noise is highly time-variant depending upon road surfacing, atmospheric conditions, age and state of components, speed of motion, and engine condition, to name just a few uncontrollable factors. Further, most brake squeal is produced by vibration of the brake components, especially the pads and discs, due to resonance. Thus, the source of the squeal is distributed over the disk, pad and calliper surfaces, giving it a significant spatial spread. The squeal is narrow-band in nature, with the squeal frequencies typically ranging from 1 to 16 kHz. Also, the squeal frequencies are usually different for each wheel, and dependent upon the speed of the automobile when braking, age of the components, environmental conditions (e.g., heat, moisture), brake pressure, etc. Analysis and quantification of the various factors inducing brake squeal would therefore help the manufacturers further reduce this contributor to the acoustic footprint.

Depending upon the aim of the analysis, this can be conducted from several different perspectives. One good classification, suggested in Mauer and Haverkamp (2007), consists of the following three approaches:

(a) A *driver-focussed* analysis, which considers the perceivable attributes of brake squeal in the interior of the automobile, with the aim of optimising brake noise and vibration as part of the driving experience.
(b) A *pedestrian-focussed* analysis, which considers the perception of brake noise *exterior* to the automobile to minimise the exterior radiation of noise.
(c) A *component-focussed* analysis, where the assessment is focussed on the source of the noise, in order to apply measures of noise reduction at the source.

[a)]Part of the work was done during the respective Ph.D. theses of the first and second authors at the Ruhr-Universität Bochum, Germany.
[b)]Electronic mail: nilesh.madhu@ugent.be

Whereas significant research has been conducted towards the perceptual analysis of brake squeal with respect to the first two assessment goals (see, e.g., Attia *et al*., 2006; Philippen *et al*., 2010), work is still required for achieving the third goal—that of component-based analysis. Such an analysis is beneficial in that the data are more appropriate to understand the excitation process leading to brake squeal and could, thereby, allow for the development and application of suitable noise reduction measures *at the source* (see, e.g., Hou *et al*., 2009; Liang and Yamaura, 2014; Lü *et al*., 2017a,b; Pan *et al*., 2019).

A first step in this direction is the *detection* of brake squeal and the *localisation* of this squeal to the contributing brake(s). This information can then be combined along with the other local parameters such as temperature, brake-pedal pressure, etc., in the subsequent analysis. Therefore, from the acoustic signal processing point of view, we have the following:

(1) An environment that contains several narrowband sources, the *a priori* knowledge of whose approximate positions are available (the four wheel positions).
(2) At any time none, one, or several of these sources could be "active" (i.e., emit a squeal), occupying possibly overlapping frequencies. The presence of active sources within an observed time range constitutes what we shall define as an acoustic event.

Brake squeal localisation defined in this context consists of detecting such acoustic events and locating them to their contributing sources (generator brakes). For each detected squeal frequency, given that there are four brakes, there are $(2^4 - 1)$ possible combinations[1] that could have contributed to the generation of the squeal. Thus, selecting the right combination constitutes the solution to the localisation problem. In other words, brake squeal localisation may be seen as a multiple-hypothesis testing or a *detection* problem based on the directional properties of the sound field received at appropriately situated sensors.

While the intuitive solution for brake squeal detection and localisation would be to analyse signals from individual microphones mounted next to the sound sources, i.e., next to the brake in the wheel casing, or from accelorometers placed on the brake itself, this setup suffers from several disadvantages that make the implementation impractical. First, the wheel casings are not easily accessible, making installation difficult and requiring special facilities. Furthermore, this region is prone to high temperatures and often collects dust and dirt. These could affect the sensors and could lead to a more-or-less complete system breakdown due to the lack of sensor-redundancy. This is compounded by the fact that repair or replacement in the field is difficult. Performing brake squeal localisation using single microphones in the wheel casing requires a comparison of the amplitudes or powers of the recorded signals. To make the right decision, therefore, the microphones must be perfectly calibrated, requiring expensive measurement microphones. Even under perfect calibration, dust or dirt sticking to any one sensor will change its frequency response and replacing the defective sensor would require recalibration of the system.

Additionally, brake squeal from any one brake may lead to almost identical amplitudes at all the microphones. Thus, a distinction based on amplitudes is not easy. Furthermore, accelerometers mounted near the brake pads are difficult to secure in position mechanically, due the heat and vibration.

In Madhu *et al*. (2014), a microphone array based solution was proposed as an alternative. This approach exploited the directional properties of the sound field; hence, the arrays were not constrained to be in close vicinity of the brakes and were mounted under the chassis, which is easily accessible. Furthermore, it was argued in Madhu *et al*. (2014) that techniques based on sound field directivity are relatively robust against gain mismatches amongst the microphones, so an accurate calibration was not required. The redundancy of microphones in the array, combined with an appropriate algorithm to detect defective sensors (see, e.g., Madhu and Martin, 2008b, and references therein) not only prevents a system breakdown but additionally allows correct localisation, taking the changed array geometry into account. If required, an entire array can be replaced *in the field* since the arrays are easily accessible. This makes maintenance and servicing of the system easy.

It was further concluded in Madhu *et al*. (2014) that, given the narrowband nature of brake-squeal, the well-known broadband approaches (e.g., Benesty, 2000; Chen *et al*., 2003; Knapp and Carter, 1976; Madhu and Martin, 2005; Roth, 1971; Scheuing and Yang, 2006; Talantzis *et al*., 2005; Yoon *et al*., 2006) should not be used to localize the squeal to the generating brakes and that algorithms that lend themselves to solving this problem should be based on narrowband, approaches for each squeal frequency (cf. DiBiase *et al*., 2001; Paulraj and Kailath, 1986; Schmidt, 1981; Thiergart *et al*., 2016; Wax, 1992; Wax and Kailath, 1984). Since such narrowband approaches are implicitly based on the *phase* of the pairwise cross-power-spectra of the microphone signals, a weighted variant of the narrowband steered response power phase transform (SRP-PHAT) algorithm was adopted in Madhu *et al*. (2014) and an approach based on *ad hoc* formulation of likelihoods was developed. Here, we propose an alternative method, based on more solid theoretical foundations, that can systematically account for the different uncertainties. We further demonstrate the improved consistency of this approach over the existing state-of-the-art in a real setting.

The remainder of the paper is organised as follows: we first introduce the signal model of Madhu *et al*. (2014) in Sec. II, followed by the system overview in Sec. III (array placement and hypothesis nomenclature). In Sec. IV, we briefly describe the algorithm of Madhu *et al*. (2014). In Sec. V, we detail the proposed approach. The developed algorithm is next tested on both simulations and data obtained under realistic conditions, and its performance relative to the state-of-the-art summarised.

II. SIGNAL MODEL

Since we deal with the narrowband localisation of the sources, we first transform each microphone signal into the short-time Fourier domain. This frame-wise spectral

J. Acoust. Soc. Am. **146** (6), December 2019

Madhu *et al*.   4899

representation of a discrete-time signal $x(n)$ is obtained by a K-point discrete Fourier transform (DFT) on overlapped, windowed signal segments. The corresponding representation of the signal is $X(k,b)$, where k is the discrete frequency bin index and b is the frame index.

Consider an array of M microphones at positions $\mathbf{r}_m = (x_m, y_m, z_m)^T$ capturing the signals emitted from Q sources located at $\mathbf{r}_q = (x_q, y_q, z_q)^T$. The signal at microphone m may be *approximated* as (Madhu and Martin, 2008a)

$$X_m(k,b) \approx \sum_{q=1}^{Q} A_{0,mq}(k)S_{0q}(k,b) + V_m(k,b), \qquad (1)$$

where $A_{0,mq}(k)$ is the transfer function from $\mathbf{r}_q$ to $\mathbf{r}_m$; $S_{0,q}(k,b)$ is the qth source signal and $V_m(k,b)$ is the noise component at microphone m. The approximation is a result of truncating the support of the signals to finite length. Each $A_{0,mq}(k)$ is further represented as $A_{0,mq}(k) = |A'_{0,mq}(k)|e^{-j\Omega_k \tau_{mq}} + A''_{0,mq}(k)$ where $|A'_{0,mq}|$ represents the gain along the direct path and $A''_{0,mq} \in \mathbb{C}$ indicates the net gain and phase smearing caused by the reflections along the indirect paths. τ_{mq} represents the *absolute* time delay of the signal from source q to microphone m along the direct path with $\Omega_k = 2\pi k f_s/K$ being the kth discrete frequency and f_s the sampling frequency.

Whereas the direct path components $A'_{0,mq} = |A'_{0,mq}(k)| e^{-j\Omega_k \tau_{mq}}$ of the transfer function are directly related to the geometric arrangement of the sources and the sensors and are key to the localisation problem, the indirect path components ($A''_{0,mq}$) are treated as disturbances. Further, dominance of the direct path is assumed (i.e., $|A'_{0,mq}| \gg |A''_{0,mq}|$), allowing us to lump the indirect component contributions with the background noise.

The model is further simplified in terms of the *relative transfer functions* (Gannot *et al.*, 2001) by considering the signals received at the first microphone through the direct path as the reference,

$$S_q(k,b) = |A'_{0,1q}(\Omega)|e^{-j\Omega_k \tau_{1q}} S_{0q}(k,b). \qquad (2)$$

Under these simplifications and stacking, the microphone signals into a vector for each time-frequency point (k,b) leads us to the following compact vector representation of the model as

$$\mathbf{X}(k,b) \triangleq \mathbf{A}(k)\mathbf{S}(k,b) + \mathbf{V}(k,b), \qquad (3)$$

with

$$\mathbf{A}(k) = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ \left|\frac{A'_{0,M1}(k)}{A'_{0,11}(k)}\right|e^{j\Omega_k \Delta\tau_{M1}} & \cdots & \left|\frac{A'_{0,MQ}(k)}{A'_{0,1Q}(k)}\right|e^{j\Omega_k \Delta\tau_{MQ}} \end{pmatrix},$$

where $\Delta\tau_{mq} = \tau_{1q} - \tau_{mq}$ is the *relative* time delay or *time delay of arrival* (TDOA) of the wavefront of source q at microphone m, with respect to the first microphone.

III. SYSTEM OVERVIEW

A. System overview

The system proposed in Madhu *et al.* (2014) consisted of four linear microphone sub-arrays, each consisting of eight elements. The inter-microphone spacing for each sub-array was chosen in order to ensure that an unambiguous decision at each frequency would be possible in the desired range and that the spatial aliasing problem (van Trees, 2002) would not occur. This led to the harmonically nested array depicted in Fig. 1. Such harmonic nesting additionally has the benefit of frequency invariant beamwidth (Ward *et al.*, 2001) in the SRP-PHAT cost function over the frequency range of interest. The four linear sub-arrays are placed along the outside border of the auto-chassis as depicted in Fig. 2. This placement preserves the dominance of the direct path from a brake to the arrays to which it has a *direct* line of sight (LOS). To clarify: array 2 can distinguish accurately between events originating front right and front left, but an event at the rear right/left brakes is not accurately localised (construction of the chassis is such that line-of-sight is obstructed). The reverse holds true for array 4. Array 1, on the other hand, cannot distinguish between an event at *front* right and *front* left. Controlled tests on a stationary vehicle demonstrated that even when the squeal event was from front left, array 1 indicated maximum spatially coherent energy from the direction of front right. The reverse is true for array 3. We infer that this behaviour is due to the shadowing of the brake at front right by the wheel, and the presence of a parallel reflecting surface in the vicinity of the brake at front left, leading to a coherent first-order reflection. Thus, array 1 perceives a source emanating from front *right* as coming from front *left*. The same holds for squeals originating from the rear wheels, while the reverse argument holds for array 3. Consequently, arrays 1 and 3 are used to distinguish reliably between events coming from the front and the rear. This discrimination capability for each array is also illustrated in Fig. 2. Accordingly, we require separate arrays for the front and rear wheels, to perform a left-right distinction and two arrays on the side to perform the front-back distinction. The reasoning for the redundancy in the front-back distinction is explained further in the following section.

B. Hypothesis nomenclature and formulation

Assume, for now, the availability of an initial stage that can detect the occurrence of brake squeal and extract the corresponding squeal frequencies. Given these frequencies at which we have an acoustic event and the directional statistics available from the arrays, the hypothesis testing approach requires us to decide which of the $2^4 - 1$ combinations of brakes generated the event. Since no single array can localise
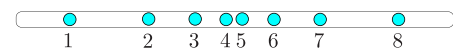

FIG. 1. (Color online) The designed array for brake squeal localisation. The spacings are symmetric about the center, with $d_{12} = 4$ cm, $d_{23} = 2$ cm, $d_{34} = 1$ cm, and $d_{45} = 0.5$ cm.
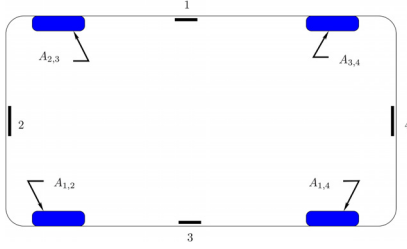
FIG. 2. (Color online) Arrangement of arrays to achieve LOS of two arrays for each brake. The subscripts for each wheel indicate the arrays having a direct LOS to that wheel. (a) Search regions for each wheel. (b) Candidate positions for array 3.

all sources, hypothesis tests which consider the complete 4-array system were not optimal. Therefore, in the hierarchical approach we proposed previously, the localisation of a squeal event to a brake was done in two steps. In the first stage, arrays 1 and 3 were used to test

(1) $\mathcal{H}_F$: squeal emanated from the front.
(2) $\mathcal{H}_B$: squeal emanated from the back (rear).
(3) $\mathcal{H}_{FB}$: squeal emanated simultaneously from the front and rear.

Depending upon the outcome of this stage, arrays 2 and/or 4 were considered to detect if the event originated from the right (R) or the left (L), i.e., if $\mathcal{H}_F$ was true, we used array 2 to test pick from the subsequent hypotheses:

(1) $\mathcal{H}_{L|F}$: squeal emanated from the brake on front left.
(2) $\mathcal{H}_{R|F}$: squeal emanated from the brake on front right.
(3) $\mathcal{H}_{LR|F}$: squeal emanated from front left and front right.

Similarly, given the occurrence of squeal from the rear, array 4 was used to decide between $\mathcal{H}_{L|B}$, $\mathcal{H}_{R|B}$, and $\mathcal{H}_{LR|B}$. Note that the accuracy of the front/back detection is critical as the next stage depends on the outcome of this test. It is to increase this accuracy and to provide robustness against sensor defects that we built in redundancy by using two arrays for this stage. The truth of the hypotheses themselves depends on the particular use of the spatial statistics.

IV. HIERARCHICAL APPROACH SUMMARY

The approach of Madhu *et al.* (2014) was based on heuristic conditions with the aim of assigning a measure of *contribution* of each brake to each acoustic event, thereby arriving at a "soft" decision. Under assumptions of independence between the squeal event at each brake, a heuristic *likelihood* $\mathcal{L}$ of activity for each brake was computed in a factorable form as in Eq. (4) below,

$$\begin{aligned}
\mathcal{L}(F,R) &= \mathcal{L}(R|F) \cdot \mathcal{L}(F), \\
\mathcal{L}(F,L) &= \mathcal{L}(L|F) \cdot \mathcal{L}(F), \\
\mathcal{L}(B,R) &= \mathcal{L}(R|B) \cdot \mathcal{L}(B), \\
\mathcal{L}(B,L) &= \mathcal{L}(L|B) \cdot \mathcal{L}(B),
\end{aligned} \tag{4}$$

where, e.g., $\mathcal{L}(F,R)$ denotes the likelihood of activity of the front, right brake; $\mathcal{L}(F)$ denotes the likelihood of activity from the front brakes; $\mathcal{L}(R|F)$ denotes the likelihood of activity of the front right brake *conditional* upon the hypothesis that the event originated from the front; and so on. The

likelihood was based on a pair-wise coherence-weighted sum of the cross-power spectrum components of the SRP-PHAT cost-function $\mathcal{J}(\mathbf{r}, k, i)$. In this notation, the dependence on frequency k, the candidate position $\mathbf{r}$ and the array i is explicit. For convenience, this function is reproduced below for a generic array as

$$\mathcal{J}(\mathbf{r}, k) = \sum_{\substack{m,m' m \neq m' \\ m,m'}} \left(e^{j\Omega_k(\Delta\tau_m(\mathbf{r}) - \Delta\tau_{m'}(\mathbf{r}))} \right. \\ \left. \cdot \frac{\Psi_{X_m X_{m'}}(k)}{|\Psi_{X_m X_{m'}}(k)|} |\Gamma_{mm'}(k)|^2 \right), \tag{5}$$

where, in accordance with the model in Sec. II, $\Delta\tau_m(\mathbf{r})$ is the TDOA that would be generated across the array, if a source were present at location $\mathbf{r}$; $\Psi_{X_m X_{m'}}(k) = X_m(k)X_{m'}^*(k)$ and $\Gamma_{mm'}(k)$ is the coherence between microphones m and m' at frequency Ω_k.

To account for the small deviations in the positioning of the microphone arrays and the spatially spread nature of the acoustic sources, the candidate positions for each brake are extended to encompass the possible spread of the source and the deviations in the microphone array, leading to a spatially averaged estimate of $\mathcal{J}$ over a candidate *region* $\mathcal{R}$,

$$\mathcal{J}(\mathcal{R}, k, i) = \int_{\mathbf{r} \in \mathcal{R}} \mathcal{J}(\mathbf{r}, k, i) d\mathbf{r}. \tag{6}$$

Since the squeal event can be reasonably assumed to be concentrated in the region around the respective wheels, we define the extent of the candidate regions as depicted by the shading in Fig. 3(a) for all the wheels and in Fig. 3(b) for array 3 in particular.

With these extended regions and the corresponding spatially averaged metrics, the marginals and the conditional probabilities were defined as

$$\begin{aligned}
\mathcal{L}(F, k) &\triangleq \frac{\mathcal{J}_F(k)}{\max(\mathcal{J}_F(k), \mathcal{J}_B(k))}, \\
\mathcal{L}(B, k) &\triangleq \frac{\mathcal{J}_B(k)}{\max(\mathcal{J}_F(k), \mathcal{J}_B(k))},
\end{aligned} \tag{7}$$

and

$$\begin{aligned}
\mathcal{L}(L|F, k) &\triangleq \frac{\mathcal{J}(\mathcal{R}_{FL}, k, 2)}{\max(\mathcal{J}(\mathcal{R}_{FR}, k, 2), \mathcal{J}(\mathcal{R}_{FL}, k, 2))}, \\
\mathcal{L}(R|F, k) &\triangleq \frac{\mathcal{J}(\mathcal{R}_{FR}, k, 2)}{\max(\mathcal{J}(\mathcal{R}_{FR}, k, 2), \mathcal{J}(\mathcal{R}_{FL}, k, 2))}, \\
\mathcal{L}(L|B, k) &\triangleq \frac{\mathcal{J}(\mathcal{R}_{BL}, k, 4)}{\max(\mathcal{J}(\mathcal{R}_{BR}, k, 4), \mathcal{J}(\mathcal{R}_{BL}, k, 4))}, \\
\mathcal{L}(R|B, k) &\triangleq \frac{\mathcal{J}(\mathcal{R}_{BR}, k, 4)}{\max(\mathcal{J}(\mathcal{R}_{BR}, k, 4), \mathcal{J}(\mathcal{R}_{BL}, k, 4))},
\end{aligned} \tag{8}$$

where

$$\begin{aligned}
\mathcal{J}_F(k) &= \mathcal{J}(\mathcal{R}_{FL}, k, 1) + \mathcal{J}(\mathcal{R}_{FR}, k, 3), \\
\mathcal{J}_B(k) &= \mathcal{J}(\mathcal{R}_{BL}, k, 1) + \mathcal{J}(\mathcal{R}_{BR}, k, 3).
\end{aligned} \tag{9}$$

J. Acoust. Soc. Am. **146** (6), December 2019

Madhu *et al.*   4901

(a) Search regions for each wheel
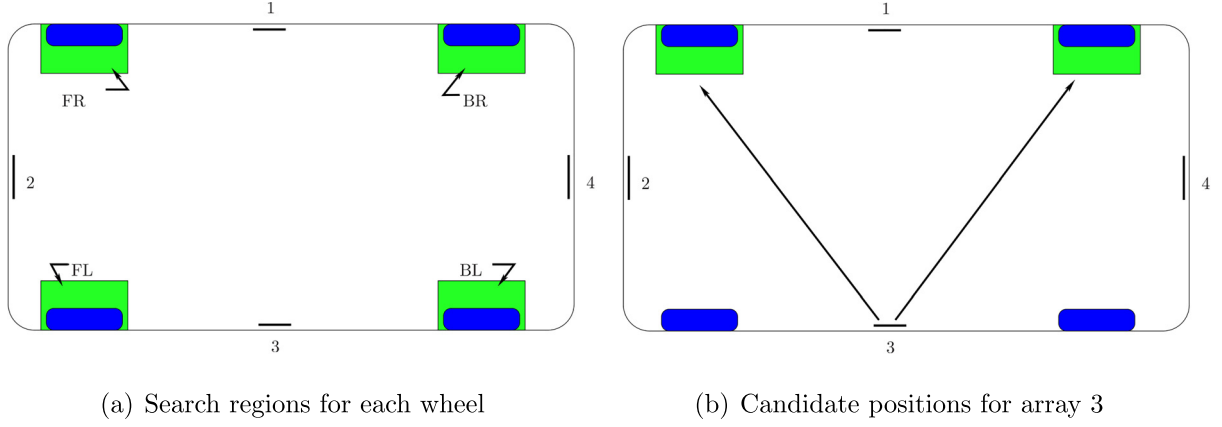
(b) Candidate positions for array 3

FIG. 3. (Color online) Candidate search regions. The shaded regions indicate the *a priori* knowledge of regions of source presence, which is used to obtain the candidate search regions for the cost function computation as in Eq. (6). The search region for array 3 is, further, illustrated for convenience.

This gives us a normalised value for the occurrence of a squeal event—akin to the *probability* of that event. At this stage, we may take a hard decision regarding the activity of a particular brake, which is done by thresholding the likelihood values obtained from Eq. (4) by substituting the values of Eqs. (7) and (8). This threshold, $\Upsilon_{\mathcal{L}}$, is empirically set.

V. PROPOSED APPROACH SUMMARY

This approach is derived from the sequential hypothesis testing framework based on maximum likelihood (Kraus, 1993; Maiwald, 1995). However, these approaches cannot be used directly since they require accurate knowledge of the signals and the propagation vector and such information is hard to gather for brake squeal localisation. The reason lies in part in the insufficient knowledge regarding the source locations (the positions are only approximately known, the spatial extent during a squeal is unknown and time and frequency variant. Estimating these parameters during the localisation procedure is also difficult, since the squeal is only present for a short time interval). In addition, we require a model for the signal distortion caused by the acoustic environment under the automobile. This is further complicated by the fact that such a model is time variant and depends not only on the relative source–receiver positions, but also on the external environment, which is continually changing. Thus, given the short time frames in which the detection and localisation must be done, applying the ML framework requires a high-dimensional optimisation, which increases the computational complexity and makes a real-time implementation impossible. This is the motivation to find an approach that is not only real-time capable (low computational complexity) but also more tolerant and robust against such imperfect knowledge. In this section, we shall develop one such method.

Since each array principally evaluates two regions, consider one such generic array, with the corresponding sources $S_1(k,b)$ and $S_2(k,b)$. We further make the simplifying assumption that the sources are *spectrally disjoint*, i.e., each active frequency is dominated by one source, which we require to localise. This assumption is realistic since the

squeal frequencies and the presence of squeal depend on several factors (hydraulic pressure, humidity, dust, state of brake shoe and disk, speed, etc.), which will be different for any two brakes. Note this contrast to the previous algorithm, which allows multiple wheels to contribute to a squeal event at a given frequency and time frame. The spectral disjointness assumption further implies that, given an active frequency, it *must* be allocated to *one* source. To do this, we sequentially test each of the two regions for the presence of source activity whilst *blocking* out contributions from the other region (treated as a *clutter* region). The approximate knowledge of the source positions can be used for designing such blocking systems with sufficient tolerance to compensate for possible inaccuracies in the model. Assume for now that we have such a blocking system $\mathbf{P}_q^\perp(k)$ for region 1 at frequency bin k. Then, define $\mathbf{Y}^{(1)}(k,b)$ as

$$\mathbf{Y}^{(1)}(k,b) = \mathbf{P}_1^\perp(k)\mathbf{X}(k,b)$$
$$= \mathbf{P}_1^\perp(k)\left(\sum_{q=1}^2 \check{\mathbf{A}}_q(k)S_q(k,b) + \mathbf{V}(k,b)\right)$$
$$\approx \mathbf{P}_1^\perp(k)\big(\check{\mathbf{A}}_2(k)S_2(k,b) + \mathbf{V}(k,b)\big), \qquad (10)$$

where $\check{\mathbf{A}}$ indicates the relative propagation vector from source to receiver [i.e., the signal model from Eq. (3)], but taking into account the unknown disturbances. The approximation in the last step holds as, by designing the blocking system to have a suitable tolerance, we may negate the contributions from the blocked source 1. Similarly, we may negate the contribution of source 2 to obtain: $\mathbf{Y}^{(2)}(k,b)$. The use of such a blocking system to obtain the resultant signals $\mathbf{Y}^{(q)}(k,b)$ is key to our approach, which we term the *maximum choice* (MaxChoice) algorithm. Note that such a blocking matrix is well-known in the beamforming community for generating the noise estimate in, e.g., adaptive beamforming approaches such as the *generalised sidelobe canceller* (Gannot *et al.*, 2017; Griffiths and Jim, 1982).

We next compute the residual energy in $\mathbf{Y}^{(1)}(k,b)$ and $\mathbf{Y}^{(2)}(k,b)$, over the B time records. This may be compactly expressed as, when source q is blocked, as

$$\breve{\Psi}_{qq}(k) = \sum_{b=1}^{B} \mathrm{tr}\left(\mathbf{Y}^{(q)}(k,b)\mathbf{Y}^{(q),H}(k,b)\right)$$

$$\approx \mathbf{P}_q^{\perp}(k)\left(\sum_{b=1}^{B}\left(\breve{\mathbf{A}}_{3-q}(k)\breve{\mathbf{A}}_{3-q}^{H}(k)|S_{3-q}(k,b)|^2\right.\right.$$

$$\left.\left. + \mathbf{V}(k,b)\mathbf{V}^{H}(k,b)\right)\right)\mathbf{P}_q^{\perp,H}(k). \tag{11}$$

Recognise that this is a measure of the energy received by the microphone from spatial regions *other* than the region associated with the presence of source q. Once this residual energy has been computed, a binary decision is made for that bin as follows:

$$\frac{\breve{\Psi}_{22}(k)}{\breve{\Psi}_{11}(k)} \underset{\mathcal{H}_2}{\overset{\mathcal{H}_1}{\gtrless}} 1, \tag{12}$$

with $\mathcal{H}_q$ representing the hypothesis that the source originated from region q. Note that in this approach we do not perform explicit source localisation. Rather, by configuring the blocking system with suitable tolerance around the *nominal* source positions, we compensate for imperfections in the knowledge of the propagation model and source position.

When the propagation model is well-known and the source positions are available, it is straightforward to extend this approach to perform a hypothesis test for the presence or absence of source activity *independently* in each region. Thereby, it is also possible to handle the case of correlated or even *coherent* sources, and the assumption of spectral disjointness is not necessary. Indeed, one may consider such a system to be a generalised form of the approach presented in Tadaion *et al.* (2007), that better uses the *a priori* knowledge available regarding the nominal source positions and spatial extent.

MaxChoice is summarised in Fig. 4.

A. Design of the blocking system

We now address the issue of design of the blocking matrix for a generic frequency k, an M channel array and a blocking region defined as $\mathcal{R}_q$. Denote by $(M \times N_q(k))$, the dimensions of the blocking matrix. The frequency-dependent parameter $N_q(k)(< M)$ is determined such that we minimise the contribution of sources from $\mathcal{R}_q$, i.e., find $\mathbf{P}_q^{\perp}(k)$ such that

$$\int_{\mathcal{R}_q}||\mathbf{P}_q^{\perp}(k)\mathbf{A}(\mathbf{r},k)||^2\mathrm{d}\mathbf{r} \approx 0$$

$$\Rightarrow \mathbf{P}_q^{\perp}(k)\left(\int_{\mathcal{R}_q}\mathbf{A}(\mathbf{r},k)\mathbf{A}^{H}(\mathbf{r},k)\mathrm{d}\mathbf{r}\right)\mathbf{P}_q^{\perp,H}(k) \approx 0. \tag{13}$$

Defining

$$\mathbf{R}_q(k) = \int_{\mathcal{R}_q}\mathbf{A}(\mathbf{r},k)\mathbf{A}^{H}(\mathbf{r},k)\,\mathrm{d}\mathbf{r}, \tag{14}$$

we see that Eq. (13) is minimised if we select $\mathbf{P}_q^{\perp}(k)$ to span the orthogonal complement space of $\mathbf{R}_q(k)$, i.e., if the eigenvectors of $\mathbf{R}_q(k)$ are denoted as $\mathbf{U}_i(k)$,

$$\mathbf{P}_q^{\perp}(k) = \left(\mathbf{U}_{M-N_q(k)+1}(k), ..., \mathbf{U}_M(k)\right)$$

$$\left(\mathbf{U}_{M-N_q(k)+1}(k), ..., \mathbf{U}_M(k)\right)^{H}, \tag{15}$$

where $N_q(k)$ is estimated, for a given threshold Υ $(0 \leq \Upsilon \leq 1)$, in the following manner:

$$N_q(k) = \underset{N}{\mathrm{argmin}}\left|\Upsilon - \frac{\sum_{i=1}^{M-N}\lambda_i(k)}{\sum_{i=1}^{M}\lambda_i(k)}\right|, \tag{16}$$

where the $\lambda_i(k)$ are the eigenvalues of $\mathbf{R}_q(k)$, sorted in the descending order of magnitude. In effect, Eq. (16) selects the

```
1 Initialisation: ∀ arrays i do

   (a) Define appropriate blocking regions R_{i,q}, q = {1,2} from Table I

   (b) ∀ q,k define corresponding blocking matrices P_{i,q}^⊥(k) using (15) &
       (16)

2 Analysis: ∀ active frequencies k do

   (a) Use data of arrays 1 & 3 and (12) to decide if the squeal at
       frequency k comes from:

           H_F: the front (F) or
           H_B: the back (B)

   (b) If H_F, use data of array 2 and (12) to decide between
       hypotheses:

           H_{FR}: the front right (FR) or
           H_{FL}: the front left (FL)

   (c) Else if H_B, use data of array 4 and (12) to decide between
       hypotheses:

           H_{BR}: the back right (BR) or
           H_{BL}: the back left (BL)
```

FIG. 4. Summary of the MaxChoice approach.

J. Acoust. Soc. Am. **146** (6), December 2019
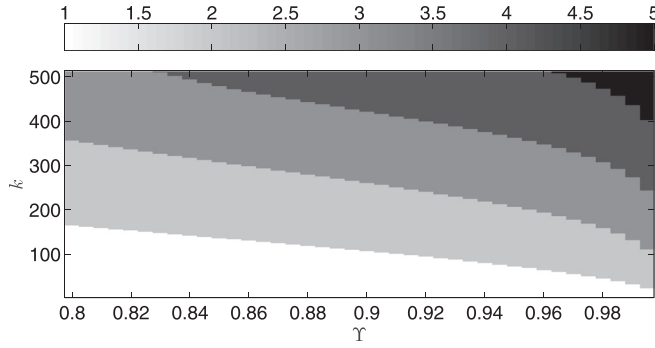
Madhu *et al.* 4903

FIG. 5. Target region dimensions for the array shown in Fig. 1. The sampling frequency is $f_s = 32$ kHz and a $K = 1024$ point DFT has been used. The target region dimension and the orthogonal complement space dimension add up to M for each combination of (k, Υ). (a) $\Upsilon = 0.95$; (b) $\Upsilon = 0.99$.

dimension of the orthogonal complement space by setting a threshold on the energy that is present in the target region. The dimension of the target region subspace and the orthogonal complement space are illustrated below for an azimuthal blocking region $\mathcal{R}_\theta = [25°, 65°]$, for various values of Υ.

As may be seen in Fig. 5, the target region dimensions are low for the lower frequencies. This is because of the lack of directivity in these frequencies, making $\mathbf{A}(\mathbf{r}, k)$ in Eq. (13) almost independent of $\mathbf{r}$. As k increases, so does the directivity, and the matrix $\mathbf{R}_q(k)$ increases in rank. The blocking capabilities of such a designed system is illustrated in Fig. 6 for two sample values of Υ.

The blocking matrix is computed according to (14)–(16), where the integration is either performed numerically or, under simplifying assumptions, may also be obtained in a closed form solution (see Appendix A). Note that the broad spread of the blocking region ($\pm 20°$ about the mean source DOA) is chosen to account for the inaccuracies in array positioning, source spread and propagation model uncertainties. The target regions for the two sources corresponding to each array are defined as in Table I. Note, further, that the tolerance regions are the same for *all* arrays, requiring only a single computation of $\mathbf{P}_q^\perp(k)$, unlike the hierarchical approach where the cost function parameters for each array have to be computed separately. Further, this approach is easy to tailor to almost any automobile size and does not require a separate recalibration of the search regions as in the hierarchical approach.

TABLE I. Blocking regions for the two sources corresponding to each array.

$\mathcal{R}_{\theta_1}$	$\mathcal{R}_{\theta_2}$
$[25°, 65°]$	$[115°, 155°]$

VI. EXPERIMENTAL EVALUATION

The presented approaches to brake squeal localisation were evaluated using a VW Touran. The recordings of brake squeal were made under different driving conditions and surroundings. As mentioned previously, the signals were analyzed in $T = 200$ ms time segments. Each segment was decomposed into its short-time spectrogram using a windowed, K point DFT, with a frame shift of O samples, as depicted in Fig. 7. Such spectrograms were constructed for each channel of every array and evaluated for the presence of active frequencies. If such frequencies were found, the hypothesis tests were carried out in the corresponding frequency bins. The analysis parameters are presented in Table II. These parameters yield, for our application, an acceptable tradeoff between computational complexity, storage requirements, and efficient processing of information for instrumental and perceptual analysis. For the spectral analysis, the von Hann or the Hamming windows are usually used due to their good spectral characteristics. The threshold $\Upsilon_{\mathcal{L}} = 0.7$ was empirically set based on the experiments we did in the controlled scenario (see further). The DFT was computed using the efficient *fast* Fourier transform (FFT) technique.

A general difficulty in the evaluation stems from the absence of reliable ground truth. Even for the experienced listener, it is not easy to localise sounds while driving. Therefore, we first validated the approaches on a controlled scenario, where the automobile was at standstill and the brake squeal was simulated by manually exciting the brake discs with a shaker. For the subsequent evaluation on the realistic recordings, we base our evaluation on the plausibility of the results (we know the approximate resonance frequencies of the brakes and thereby can see if signals detected at these frequencies are correctly assigned).

A point of note here: the finite spectral resolution associated with the DFT coupled with the frequency jitter of the sources smears each narrowband acoustic event across more

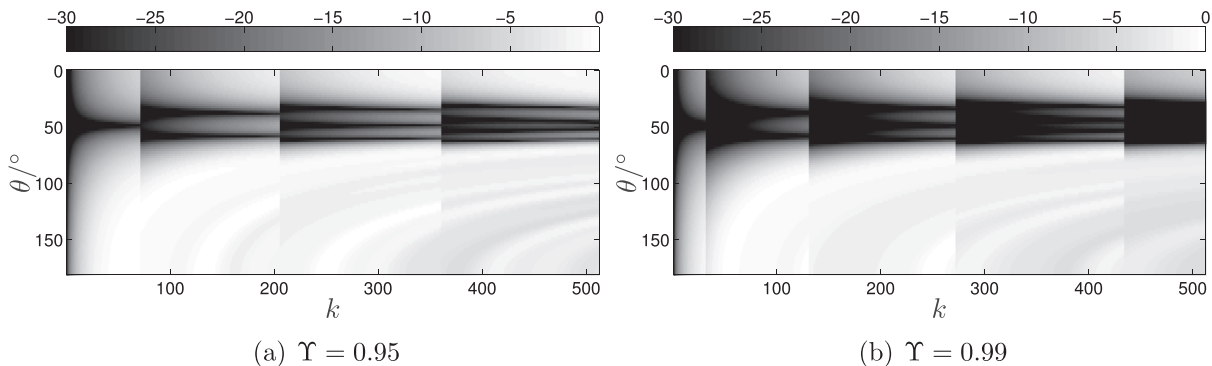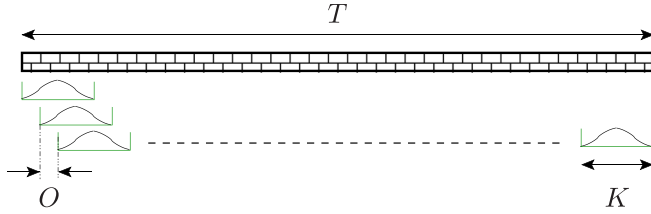(a) $\Upsilon = 0.95$

(b) $\Upsilon = 0.99$

FIG. 6. Performance of the blocking system in dB for the array of Fig. 1, for a sampling frequency of $f_s = 32$ kHz and a $K = 1024$ point DFT. The discontinuities in the plot are the frequency bins at which the dimensions of the orthogonal complement space change. The blocking region was set to $\mathcal{R}_\theta = [25°, 65°]$.

FIG. 7. (Color online) Schematic of the data analysis for a microphone signal illustrating the concept of a signal segment of length T ms, its decomposition into frames of length K with a frame shift of O between two successive frames. The fact that the data-frames are windowed before the computation of the DFT is also schematically illustrated. (a) Sample signal spectrum, (b) localisation results, hierarchical approach, (c) localisation results, MaxChoice approach.

than one frequency bin. The accretion of such adjacent frequency bins builds what we denote as a frequency *band* and the presence of an active band characterises an acoustic event.

Both the hierarchical approach and the MaxChoice approach are modified to take such spectral smearing into account. In the former case, the probability of activity is computed using the spectrally-averaged cost functions for each array and in the latter case, the decision is based on a weighted average of the decisions in the individual bins of a band. As the MaxChoice approach essentially makes one decision per band, the maximum allowable bandwidth of a band in this approach is constrained to about 100 Hz. No such restriction is placed on the bandwidth in the hierarchical case, as it is capable of making multiple decisions at any band, and thus the smearing of two different events into one band should not affect the performance of this approach.
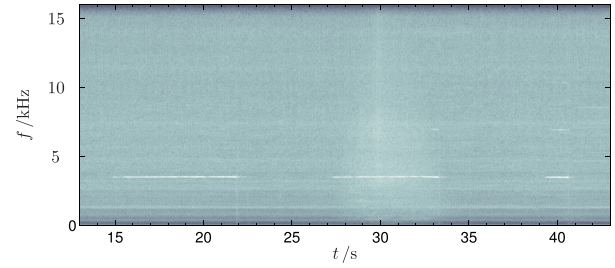
Some sample results are presented in Figs. 8–10, where one may easily discern the squeal regions in the reference spectrogram. The results obtained from the localisation approaches are depicted on this spectrogram in a colour coded manner consistent with Table III.

We begin with a localisation scenario where only a single active frequency is present. The predominantly active brake in this case was BR. Both approaches are in agreement in terms of their localisation result. At $t = 21.8$ s, the hierarchical approach performs multiple decisions whereas the MaxChoice approach makes a single decision consistent with the time history until that point. Note that in case of multiple decisions in the hierarchical approach, the results are presented for the localised brakes in the order FR, FL, BL, and BR. Thus, at $t = 21.8$ s, where the hierarchical approaches localises the squeal simultaneously to FL, BL, and BR, the results are presented by bars of the appropriate colour, slightly frequency shifted, in the above mentioned order.
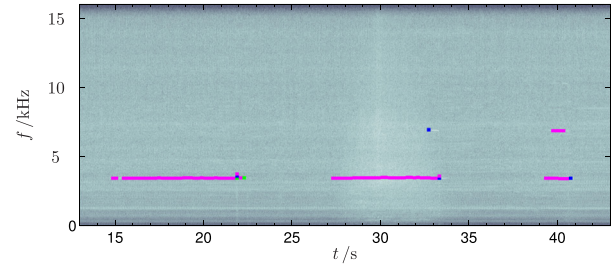
Figures 9 and 10 present a more dynamic case, where multiple brakes squeal simultaneously and at multiple, closely spaced frequencies.

(a) Sample signal spectrum

(b) localisation results, hierarchical approach

(c) localisation results, MaxChoice approach

FIG. 8. (Color online) Single active frequency, dominantly from BR. In general, both the proposed approaches are in good agreement regarding the localisation. The hierarchical approach makes multiple decisions at $t = 21.8$ s, as compared to the single decision of the MaxChoice approach. Further, the hierarchical approach detects, but does not localise the harmonic in the three time segments from 32.8 to 33.4 s. Such behaviour is examined further in Sec. VI A. This recording was made as the car was moving along a ramp in a tunnel—thus under exceedingly reverberant conditions. (a) Sample signal spectrum, (b) localisation results, hierarchical approach, (c) localisation results, MaxChoice approach.
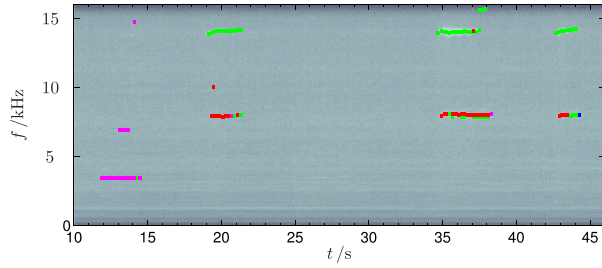
A. Relative comparison of the approaches

We now present a relative comparison of the hierarchical and the MaxChoice approach, evaluated on the available data. The comparison is according to four different criteria:

(a) Detection and localisation rate.
(b) Validity of the spectral disjointness assumption in MaxChoice.
(c) Agreement between the two approaches.
(d) Consistency of the two approaches during disagreement in the localisation result

1. Detection and localisation rate

This criterion measures the number of instances where each algorithm detects and localises a squeal. This is indicated in Table IV, where $\mathcal{Z}$ denotes the number of localisation
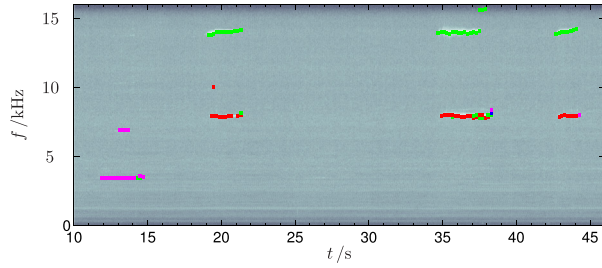
TABLE II. Parameters for the experimental evaluations.

f_s (kHz)	DFT length (ms)	Frame shift (ms)	Window/ length (ms)	Likelihood threshold $\Upsilon_{\mathcal{L}}$
32	32	8	von Hann/32	0.7

J. Acoust. Soc. Am. **146** (6), December 2019

Madhu *et al.* 4905

(a) Sample signal spectrum

(b) localisation results, hierarchical approach
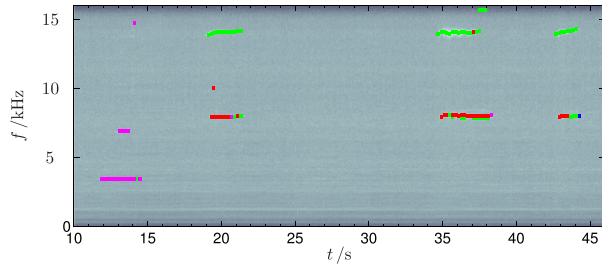
(c) localisation results, MaxChoice approach

FIG. 9. (Color online) Sample results for a more dynamic scenario. Multiple brakes squeal simultaneously and at multiple frequencies. Again, in general, both approaches are in good agreement. The period between $t = 35\,\mathrm{s} - t = 39\,\mathrm{s}$ is of interest. Two closely spaced traces are visible at $f \approx 8\,\mathrm{kHz}$. This is treated by the hierarchical approach as a *single* band, which is then localised to *both* FR and FL, and presented in the order FR, FL. The MaxChoice approach, due to its constrained maximum bandwidth, localises each trace independently, the lower to FL and the upper to FR. This explains the apparent discrepancy between the approaches. Note that the end result of localisation is essentially the same from both approaches. (a) Sample signal spectrum, (b) localisation results, hierarchical approach, (c) localisation results, MaxChoice approach.

decisions and $\bar{\mathcal{Z}}$ represents those segments where a squeal was detected but no localisation decision was taken. Additionally indicated is the number of soft decisions taken by the hierarchical approach ($\mathcal{Z}_\mathrm{s}$). We see that the data contains 1290 instances of *detected* and 895 instances (under MaxChoice) of *localised* squeal events. For any particular time frame and frequency, we say we have *detected* a squeal event if at least one array marks this frequency as active in the observed time frame. The discrepancy in the number of detected and localised events arises due to the following reasons:

C1: When the SNR is low, there are squeal events that are detected only by the one or the other array, and therefore do not contribute to the localisation result.
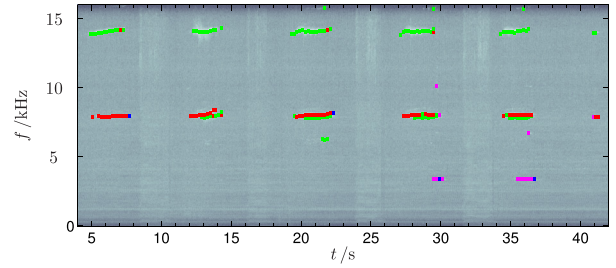
C2: In the *hierarchical* approach, given the zero upper threshold value of the cost function values, some
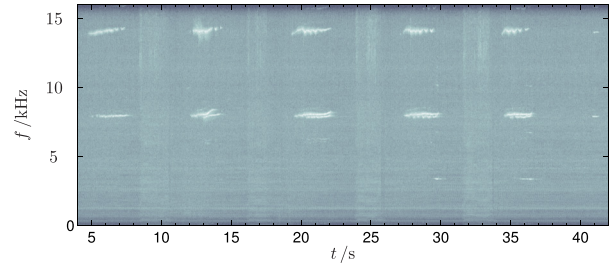
(a) Sample signal spectrum
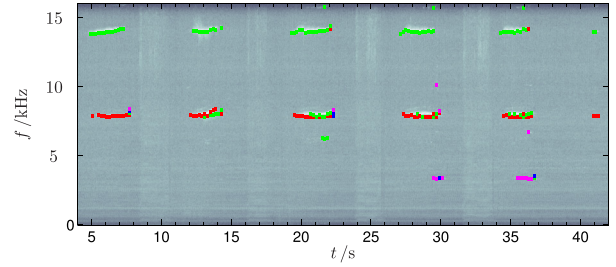
(b) localisation results, hierarchical approach

(c) localisation results, MaxChoice approach

FIG. 10. (Color online) Sample results for a multi-squeal scenario. Again, both the proposed approaches detect the squeal as being from different brakes and localise them correspondingly. Again, as in Fig. 9, the multiple traces around around $f \approx 8\,\mathrm{kHz}$ are analysed as a single band in the hierarchical approach, and localised to FR and FL, whereas the MaxChoice approach localises each trace independently. Thus, while the graphical presentation may seem different, the end result of both approaches is essentially the same. (a) Sample signal spectrum and noise floor estimate, (b) Local SNR estimate.

frequencies where a squeal is present (i.e., detected by all the arrays) may not be localised as the cost function assumes positive values over the corresponding search regions. Manual examination of such instances indicate that in these cases the minimum of the cost function is shifted to lie *outside* the demarcated search region. This could happen due to temporary divergence of the direct path as, for example, when turning the wheels. The hierarchical approach is more sensitive to such misalignments as compared to the MaxChoice approach due to the stronger influence of the cost function in the decision chain in this case, as compared

TABLE III. Colour-coding for the different wheels.

FR	FL	BL	BR
🟥	🟩	🟦	🟪

TABLE IV. Performance comparison, hierarchical and MaxChoice—Detection and localisation capability.

	$\mathcal{Z}$	$\bar{\mathcal{Z}}$	$\mathcal{Z}_s$
Hier.	861	456	27
MaxCh.	895	395	N/A

to the binary decision propagation based on relative differences in the MaxChoice case.

C3: In very rare cases, the MaxChoice approach may also fail to make a decision. This occurs when, for example, in the first stage the squeal is localised to the front of the automobile, but array 2, responsible for the subsequent FL/FR decision, does not detect the squeal at that frequency and time frame.

Since C3 is very rare (based on manual examination of the results), we take the MaxChoice approach as the baseline and assume that the lack of decision here is principally due to C1. As this condition must also hold for the hierarchical approach, the difference in the $\mathcal{Z}$ and $\bar{\mathcal{Z}}$ values between the two algorithms should indicate the frequency of occurrence of C2, a drawback of the hierarchical approach. In this context we see that the hierarchical approach fails to make a decision in an additional 61 instances, as compared to the MaxChoice approach.

2. Validity of the spectral disjointness assumption of MaxChoice

We analyze here the soft decisions taken by the *hierarchical* approach, indicating the number of times multiple wheels have been allocated to a specific event. This metric should give us an idea of the validity of our assumption regarding the temporal and spectral disjointness of the squeal signals originating from different brakes. We see from Table V, where $\mathcal{Z}_{(i)}$ indicates the number of instances where i wheels were localised, that only in 27 cases was a soft decision taken, a percentual amount of 3.4%, indicating the validity of our spectral disjointness assumption for the MaxChoice approach. Furthermore, we see that in 23 of these cases, one of the decisions of the hierarchical approach correspond to the decision taken by the MaxChoice approach. The four instances where no agreement is reached with MaxChoice are engendered by the occurrence of condition C3, which forces MaxChoice to a "no-decision" state.

3. Agreement between the two approaches

This is presented in Table VI for two cases. In the first case, we consider the instances where the squeal has been detected by all arrays and therefore must be localised at least

TABLE V. Performance comparison, hierarchical and MaxChoice—Analysis of soft decisions by the hierarchical approach.

$\mathcal{Z}_s$	$\mathcal{Z}_{(2)}$	$\mathcal{Z}_{(3)}$	$\mathcal{Z}_{(4)}$	$\mathcal{Z}_{\text{agree}}$ MaxChoice
27(3.4%)	25	1	1	23

TABLE VI. Performance comparison, hierarchical and MaxChoice—agreement.

$\mathcal{Z}_{\text{agree}}$ (decision segments)	$\bar{\mathcal{Z}}_{\text{agree}}$ (no decision segments)
800 (89.4%)	391 (99.0%)

by the MaxChoice approach. We measure, here, how often both approaches localise such events to the same wheel. In the second case, we consider the agreement between the approaches when no decision is taken for the time segment. The discrepancy between $\bar{\mathcal{Z}}$ and $\bar{\mathcal{Z}}_{\text{agree}}$ is because of the rare occurrence of C3 for the MaxChoice approach. In general, we see a very good agreement between the two proposed approaches.

4. Consistency of the results during disagreement

This criterion presents the temporal consistency of the localisation results for the cases where both approaches are not in agreement for a squeal event. In effect, we treat the localisation result for a frame $B(l)$ and frequency k as *consistent* if the same wheel has been localised at frame $B(l\pm1)$, for that same frequency (assuming, of course, that the squeal event spans more than one time frame. Consistency in this context cannot be defined for events lasting only one time frame). We see from the presented results in Table VII that the MaxChoice approach seems to be consistent for a larger proportion of the cases where the results of the approaches were different, for detected squeal events. Note, however, that the consistency measure is heuristic in the absence of ground truth, and it is difficult to ascertain if a consistent result is indeed the true result. Therefore, these results should be treated with some caution. $\mathcal{Z}_{\text{div}}$ indicates the total number of detected squeal events for which the two approaches were not in agreement and $\mathcal{Z}_{\text{consistent}}$ indicates the number of instances from $\mathcal{Z}_{\text{div}}$, where each approach is consistent.

5. Comparison summary

Both approaches are capable of detecting and localising closely spaced harmonics that originate from different wheels. This is illustrated in Fig. 9, around frequencies of 7.9–8.1 kHz. Furthermore, both show a high degree of agreement in the localisation performance, in general making the same decision when the squeal component has a high SNR. The time frames where the approaches diverge correspond mainly to segments that contain the squeal only in a small fraction of its total length, or where the SNR is low. In such cases, the temporal consistency of the results obtained by

TABLE VII. Performance comparison, hierarchical and MaxChoice—Analysis during divergent behavior.

	$\mathcal{Z}_{\text{consistent}}$		
$\mathcal{Z}_{\text{div}}$	MaxChoice	Hierarchical	Neither
99	65	21	13

MaxChoice is higher than that of the hierarchical approach. We reiterate that in the absence of a ground-truth, such a comparative evaluation cannot completely predict which algorithm is the better choice. However, the MaxChoice approach seems to be less sensitive to positioning errors and errors in the propagation model, and might well prove to be the better algorithm.

The hierarchical approach makes a soft decision in only a small percentage of the cases. In all other cases, only a single decision is made, leading us to conclude the validity of the assumption that an active frequency stems from a single source.

## VII. SQUEAL DETECTION AND IDENTIFICATION OF ACTIVE FREQUENCIES

As mentioned previously (Sec. III B), the hypothesis testing is performed only on frequency bands where brake squeal is detected. A robust detection of the squeal and the identification of the active frequencies is, therefore, critical for the proper functioning of the localisation stage. Since brake squeal consists of narrowband harmonics, the signals recorded by the microphones would consist of such harmonics superimposed upon the (relatively stationary compared to the squeal) broadband background noise. Thus, identifying spectral components that "stick-out" of the background noise would offer a good estimate of the active frequencies.

One way to do this would be to monitor the background noise level using conventional algorithms based on long-term statistics of the signal spectra (e.g., Gerkmann and Hendriks 2011; Martin, 2001), and then to consider harmonic peaks that are significantly higher than this noise floor. This necessitates maintaining a record of the required signal statistics and computing the *a posteriori* signal-to-noise ratio (SNR) (Ephraim and Malah, 1984) at each frequency and each time frame, to identify potentially active frequencies. Thus detection of brake squeal and the identification of the active frequencies are performed in a single step. This approach has two drawbacks: (1) the computational cost associated with maintaining the noise floor statistics and computing the active frequencies in *each* time-frame and (2) the memory requirements associated with maintaining records of the noise floor for computing the statistics. Since brake squeal only occurs in a small fraction of the observed time period, most of these computations are superfluous. We therefore propose to decouple the problem of detection and frequency identification: the first stage is further broken down into its two constituent subtasks: (a) *a priori* detection of squeal presence and (b) extraction of the active frequencies, if part (a) detects a squeal.

For the initial decision regarding the presence of a squeal in a signal segment it suffices to use a simpler, computationally less expensive approach. If a squeal is detected by this approach, the local SNR estimators need be run only on *that* segment for extracting the active frequencies. The algorithm used for the *a priori* detection of squeal presence is described, for completeness, in Appendix B.

We reduce the computational and memory expense further by computing the local SNRs without maintaining

noise statistics. For this, we exploit the observation that the squeal harmonics are extremely narrowband. Hence, *order statistic methods* lend themselves nicely to detecting these harmonics. Specifically, for each segment where squeal has been detected, we perform a *median* filtering across the signal periodogram. This yields an estimate of the noise floor for that time-segment, under the assumption that the noise has a smooth spectral progression. This is a reasonable assumption given that the noise is essentially road/tyre/engine noise, predominant at the low frequencies. This segment-wise noise floor estimate can be used to determine the local SNRs and, by a proper selection of the threshold, active frequency bins can be identified. This is illustrated in Fig. 11. For further robustness, the detection of squeal and computation of the active frequencies is done, for *each* array on the basis of the *aggregated* amplitude data.
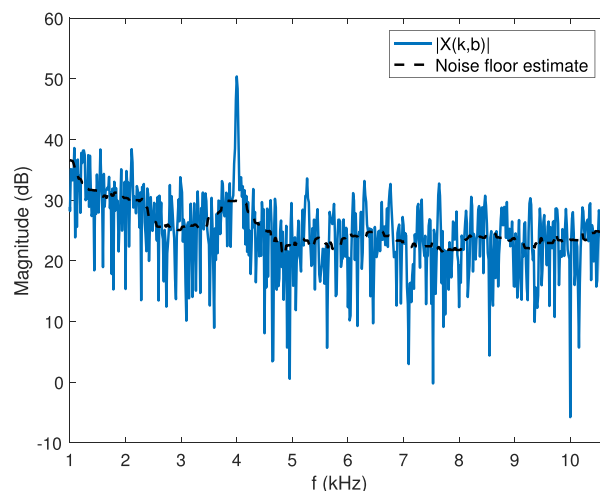
(a) Sample signal spectrum and noise floor estimate
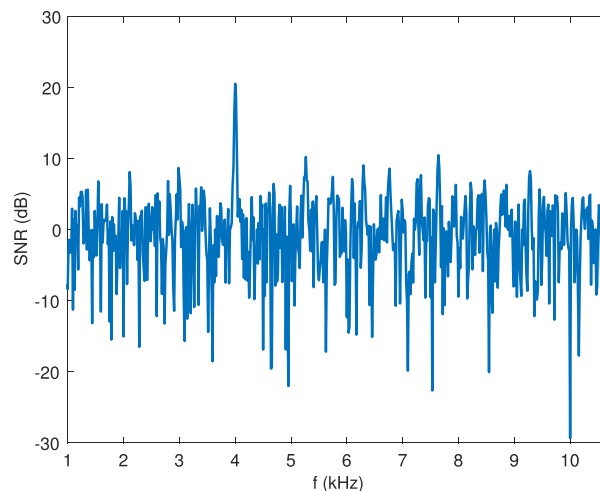
(b) Local SNR estimate

FIG. 11. (Color online) Illustration of the order-statistic method for local SNR estimation and active squeal frequency identification. Note that this approach is only carried out on signal segments where squeal is *detected*. There is no requirement to maintain signal statistics or records of past segments for computing these statistics. (a) Sub-band containing environmental noise only, (b) sub-band containing squeal.

VIII. CONCLUSIONS

This contribution has presented the design of a new localisation algorithm for the detection of brake-squeal. Given the complex and time-varying acoustic environment under the car and uncertainties in the knowledge of the source and sensor positions, localising a squeal event to the generator brakes is a challenging problem, where standard localisation approaches fail. The proposed approach builds upon the previous state-of-the-art, which treats the problem as that of a multiple-hypothesis testing problem. Under the simplifying assumptions of source disjointness, each detected squeal event (within a narrow bandwidth) is allocated to one brake. This is based on projections of the signals onto their orthogonal component subspace, where the computation of this component is designed to account for the model uncertainties. For linear arrays, the computation of the orthogonal component allows for a closed form solution. Further, for most passenger vehicle dimensions, the design of the blocking system is identical for each array, leading to a simpler architecture. Both approaches perform well on the test data. A comparison of the two approaches shows that their performance is very similar highlighting, on the one hand, the good choice of the heuristic and, on the other hand, the validity of the simplifying assumptions of the second approach. The MaxChoice approach, however, seems to present more consistent results and is computationally simpler and more versatile. More importantly, the detection approach presented here is not constrained to the problem of brake squeal localisation alone, but may be extended to other applications too. The design of robust-blocking systems, for example, may be applied to other problems where the propagation environment is not ideally known, and errors in the signal model need to be compensated for. This approach may also be applied to the hypothesis testing for the presence of *coherent* sources from *a priori* known spatial regions. Usage of appropriate blocking systems would allow for the *independent* testing of the presence or absence of source activity in each region—with a resulting lower computational complexity.

APPENDIX A: CLOSED-FORM SOLUTION FOR R

For a linear array and a blocking region demarcated in terms of the azimuth as $\mathcal{R}_\theta = [\theta_1, \theta_2]$, and under the assumption of well calibrated arrays, the propagation vector from the sources in this region to the array is defined as

$$\mathbf{A}(\theta, k) = \left(e^{-\jmath d_1 \Omega_k \cos(\theta)/c}, \ \ldots, \ e^{-\jmath d_M \Omega_k \cos(\theta)/c} \right)^T, \quad \text{(A1)}$$

with d_m being measured with respect to the center of mass of the array. The covariance matrix $\mathbf{R}(k)$ for $\mathcal{R}_\theta$ is then

$$\mathbf{R}(k) = \int_{\theta_1}^{\theta_2} \mathbf{A}(\theta, k) \mathbf{A}^H(\theta, k) \mathrm{d}\theta, \quad \text{(A2)}$$

which may be evaluated numerically, as a closed form solution to this is difficult. If we, however, make the simplifying assumption that the sources lie on the surface of a sphere with the array at the center and assume the blocking region to span the surface of the sphere between the specified angles, we may obtain a closed form solution (see, e.g., Elko, 2001) as

$$\mathbf{R}(k) = \int_0^{2\pi} \int_{\theta_1}^{\theta_2} \mathbf{A}(\theta, \varpi, k) \mathbf{A}^H(\theta, \varpi, k) \sin(\theta) \mathrm{d}\theta \mathrm{d}\varpi$$

$$= 2\pi \int_{\theta_1}^{\theta_2} \mathbf{A}(\theta, k) \mathbf{A}^H(\theta, k) \sin(\theta) \mathrm{d}\theta, \quad \text{(A3)}$$

with elements

$$\frac{R_{mm'}(k)}{2\pi}$$
$$= \begin{cases} (\cos(\theta_1) - \cos(\theta_2)), & m = m' \\ \dfrac{e^{\jmath \Omega_k d_{mm'} \cos(\theta_1)/c} - e^{\jmath \Omega_k d_{mm'} \cos(\theta_2)/c}}{\jmath \Omega_k d_{mm'}/c}, & m \neq m', \end{cases}$$
$$\text{(A4)}$$

where $d_{mm'} = d_m - d_{m'}$. Such a closed-form solution is also reminiscent of Slepian sequences or discrete prolate spheroidal sequences (Slepian, 1978), applied now to the context of spatial beampattern design.

Further, it is easy to show that if the arrays are not well calibrated in terms of amplitude, then the normalised steering vector may be reformulated as

$$\mathbf{A}(\theta, k) = \left(A_1 e^{-\jmath d_1 \Omega_k \cos(\theta)/c}, \ \ldots, \ A_M e^{-\jmath d_M \Omega_k \cos(\theta)/c} \right)^T, \quad \text{(A5)}$$

where each A_m may be modelled by a suitable probability density function with mean 1. In this case, the deterministic covariance between two elements m and m' is obtained by the statistical expectation

$$R_{mm'}^{\mathrm{det}}(k) = \mathbb{E}\left\{ R_{mm'}(k) \right\}. \quad \text{(A6)}$$

Since this expectation is not a function of the source location, it amounts to multiplying the results in Eq. (A2) by the statistical expectation of the amplitude terms: $\mathbb{E}\{A_m A_{m'}\}$. We can assume, further, these amplitude variations to be independent of each other, simplifying the computation. Under the assumption of a uniform distribution of the amplitude about $1 \pm \epsilon$, we see that the covariance matrix is unaffected for small ϵ after the expectation. On the other hand, if the perturbations in the amplitude follow other distributions, these can be factored in as well.

APPENDIX B: DETECTION OF BRAKE SQUEAL PRESENCE

The brake-squeal events are narrowband, with most of the spectral energy being concentrated in the squeal

J. Acoust. Soc. Am. **146** (6), December 2019

Madhu *et al.*   4909

frequencies,[2] giving rise to a rather well-structured spectrum in the presence of brake-squeal. The aim of the detection stage is to form an *a priori* decision on the presence of a squeal event. The prerequisites for any such approach are that it should be computationally not too demanding and yet reasonably accurate. Fast model-order estimation techniques (e.g., Cong *et al.*, 2012), based on the statistics of the eigenvectors of covariance matrices could be a possibility. If we compute the covariance matrix over small frequency bands, the presence of squeal would manifest as a dominant eigenvalue. However, this approach requires maintaining signal statistics. Further, squeal events do not always occur for long time instants, which leads to underestimation of the statistics. Hence, we propose the use of modified information theoretic tools (like entropy).

The entropy for a discrete source with an alphabet $\mathcal{A}$ and an associated probability distribution function P is defined as (Shannon and Weaver, 1949)

$$\mathfrak{H}(\mathcal{A}) = - \sum_{\forall a \in \mathcal{A}} P_a \log_2(P_a). \tag{B1}$$

It characterises the amount of disorder in the system: the entropy is maximum if all the symbols in the alphabet are equally probable and reduces as the probability distribution becomes more "peaked."

Now, consider small sub-bands of frequencies in the discrete Fourier spectrum of a signal. Let the length of each sub-band be K'. If a sub-band contains a harmonic signal, the power spectrum would be peaked at this frequency. Alternatively, if the sub-band contained only environmental noise, the distribution of power would be more or less equal along all the discrete frequency bins of that band. Similar to Renevey and Drygajlo (2001), let us define a pseudo probability distribution function along the bins k of a sub-band $\mathcal{L}$ as

$$P_{k,\mathcal{L}} = \frac{|X^{(\mathcal{L})}(k)|^2}{\sum_{k=1}^{K'} |X^{(\mathcal{L})}(k)|^2}, \tag{B2}$$

where $X^{(\mathcal{L})}(k) = X(k + (\mathcal{L} - 1)K')$ represents the DFT coefficient of bin k in sub-band $\mathcal{L}$. Such a distribution function is shown in Fig. 12.

It may be seen from Fig. 12 that the distribution defined in Eq. (B2) is a faithful representation of the underlying spectral structure. This "entropy" can then be calculated using Eqs. (B1) and (B2) and used to predict squeal presence for a fixed detection threshold $\Upsilon_{\text{apriori}}$. Note that setting the detection threshold $\Upsilon_{\text{apriori}}$ involves a trade-off between sensitivity to events at lower power and the generation of too many false alarms and depends upon the requirements of the application. If this approach detects a squeal in a signal segment, that segment is analysed in more detail and the squeal frequencies are extracted using more sophisticated algorithms.

APPENDIX C: DESCRIPTION OF HARDWARE USED

The 32-channel array with microphones and pre-amplifiers was custom-made. The microphones were Knowles FG-6163-P07 analog electret microphones with an omnidirectional characteristic and a 6 dB per octave sloping frequency response for the suppression of low frequency components at the electro-acoustic interface. Frequency response of the capsules is flat above 2 kHz.

The microphones were connected to Funk Tonstudiotechnik Type SOA-2V2 (Funk-Tonstudiotechnik, 2019) pre-amps with symmetric outputs.

For the A/D conversion and transmission to the computer, we used the hardware based on the NIST Mark III design (NIST, 2019).

A picture of the array and the casing of the data acquisition system is presented in Fig. 13 to allow the reader to get a better idea of the structure of the array and the compactness

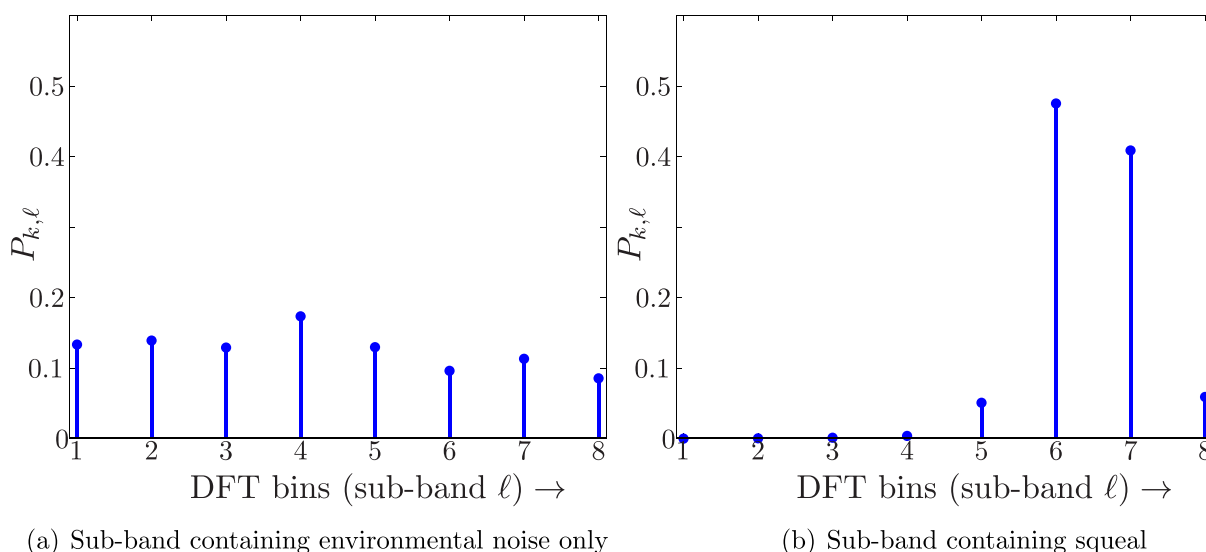(a) Sub-band containing environmental noise only

(b) Sub-band containing squeal

FIG. 12. (Color online) Probability distribution functions for cases where only noise is present (a) and where the sub-band contains an active frequency (b). Note the 'peakiness' of the second plot with respect to the almost flat curve of the first.

(a)

(b)

FIG. 13. (Color online) An illustration of the hardware used. (a) Sub-array module with protective cover removed. The microphone capsules are visible in the centre portion. All 4 arrays are similar in shape and construction. (b) Front view of compact case used for data acquisition. The preamplifiers, A/D converters, etc., are integrated into this case. The arrow indicates the ethernet connector for data transmission to the acquisition and evaluation computer.

of the data acquisition setup. Further, to enable usage in a mobile environment, the recording setup could be hooked up to a 12 V auto battery.

[1]The presence of a squeal event precludes the null hypothesis that no brakes were active.

[2]The lower frequencies containing motor noise are neglected in our considerations.

Attia, F., Zöller, T., Abdelhamid, M. K., Bretz, U., and Blaschke, P. (**2006**). "The psycho acoustical approach behind the brake squeal evaluation procedure BONI," in *Proceedings of Euro-Noise*, May 30–June 1, Tampere, Finland.

Benesty, J. (**2000**). "Adaptive eigenvalue decomposition algorithm for passive source localization," J. Acoust. Soc. Am. **107**(1), 384–391.

Chen, J., Benesty, J., and Huang, Y. (**2003**). "Robust time delay estimation exploiting redundancy among multiple microphones," IEEE Trans. Speech Audio Process. **11**(6), 549–557.

Cong, F., Nandi, A. K., He, Z., Cichocki, A., and Ristaniemi, T. (**2012**). "Fast and effective model order selection method to determine the number of sources in a linear transformation model," in *Proceedings of the European Signal Processing Conference (EUSIPCO)*, August 27–31, Bucharest, Romania, pp. 1870–1874.

DiBiase, J. H., Silverman, H. F., and Brandstein, M. S. (**2001**). "Robust localization in reverberant rooms," in *Microphone Arrays: Signal Processing Techniques and Applications*, edited by M. Brandstein and D. Ward (Springer-Verlag, Berlin), pp. 157–180.

Elko, G. (**2001**). "Spatial coherence functions for differential microphones in isotropic noise fields," in *Microphone Arrays: Signal Processing Techniques and Applications*, edited by M. Brandstein and D. Ward (Springer-Verlag, New York), pp. 61–85.

Ephraim, Y., and Malah, D. (**1984**). "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. Acoust. Speech Signal Process. **32**(6), 1109–1121.

Funk-Tonstudiotechnik (**2019**). "Funk-Tonstudiotechnik," http://www.funk-tonstudiotechnik.de/SOA-2-4.V1.pdf (Last viewed May 2019).

Gannot, S., Burshtein, D., and Weinstein, E. (**2001**). "Signal enhancement using beamforming and nonstationarity with applications to speech," IEEE Trans. Signal Process. **49**(8), 1614–1626.

Gannot, S., Vincent, E., Markovich-Golan, S., and Ozerov, A. (**2017**). "A consolidated perspective on multimicrophone speech enhancement and source separation," IEEE/ACM Trans. Audio Speech Lang. Process. **25**(4), 692–730.

Gerkmann, T., and Hendriks, R. C. (**2011**). "Noise power estimation based on the probability of speech presence," in *Proceedings of the IEEE International Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, October 16–19, New Paltz, NY, pp. 145–148.

Griffiths, L. J., and Jim, C. W. (**1982**). "An alternative approach to linearly constrained adaptive beamforming," IEEE Trans. Antennas Prop. **30**, 27–34.

Hou, J., Guo, X., Sun, G., and Zhang, J. (**2009**). "Suppression of brake squeal noise applying viscoelastic damping insulator," in *Proceedings of the 2009 International Joint Conference on Computational Sciences and Optimization*, April 24–26, Sanya, China, pp. 167–171.

Knapp, C. H., and Carter, G. C. (**1976**). "The generalized correlation method for the estimation of time delay," IEEE Trans. Acoust. Speech Signal Process. **24**(4), 320–327.

Kraus, D. (**1993**). "Approximate maximum-likelihood-schätzung und verwandte verfahren zur ortung und signalschätzung mit sensorgruppen" ("Approximate maximum-likelihood estimation and related approaches for localisation and signal estimation with sensor arrays"), Ph.D. thesis, Ruhr-Universität Bochum, Germany.

Liang, Y., and Yamaura, H. (**2014**). "Improved ensemble empirical mode decomposition method for investigating disc brake squeal signals," in *Proceedings of the 2014 12th International Conference on Signal Processing (ICSP)*, October 19–23, Hangzhou, China, pp. 205–209.

Lü, H., Shangguan, W.-B., and Yu, D. (**2017a**). "An imprecise probability approach for squeal instability analysis based on evidence theory," J. Sound Vib. **387**, 96–113.

Lü, H., Shangguan, W.-B., and Yu, D. (**2017b**). "A unified approach for squeal instability analysis of disc brakes with two types of random-fuzzy uncertainties," Mech. Syst. Signal Process. **93**, 281–298.

Madhu, N., and Martin, R. (**2005**). "Robust speaker localization through adaptive weighted pair TDOA (AWEPAT) estimation," in *Proceedings of the International Conference on Speech Communication and Technology (INTERSPEECH)*, September 4–8, Lisbon, Portugal.

Madhu, N., and Martin, R. (**2008a**). "Acoustic source localization with microphone arrays," in *Advances in Digital Speech Transmission*, edited by R. Martin, U. Heute, and C. Antweiler (John Wiley & Sons, Ltd., New York), pp. 135–170.

Madhu, N., and Martin, R. (**2008b**). "SCORE: A low complexity, robust algorithm for the detection of corrupt sensors and self-calibration of microphone arrays," in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, September 14–17, Seattle, WA.

Madhu, N., Martin, R., Rehn, H. W., Gergen, S., and Fischer, A. (**2014**). "A hierarchical approach for the online, on-board detection and localisation of brake squeal using microphone arrays," in *Proceedings of the 2014 4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA)*, May 12–14, Nancy, France, pp. 132–136.

Maiwald, D. (**1995**). "Breitbandverfahren zur signalentdeckung und -ortung mit sensorgruppen in seismik- und sonaranwendungen" ("Broadband approaches for signal detection and localisation with sensor arrays for seismic and sonar applications"), Ph.D. thesis, Ruhr-Universität Bochum, Germany.

Martin, R. (**2001**). "Noise power spectral density estimation based on optimal smoothing and minimum statistics," IEEE Trans. Speech Audio Process. **9**(5), 504–512.

Mauer, G., and Haverkamp, M. (**2007**). "Measurement and assment of noise caused by vehicle brake systems," in *Proceedings of Inter-Noise*, August 28–31, Istanbul, Turkey.

NIST (**2019**). "NIST-Mark III microphone array design details," https://www.nist.gov/sites/default/files/documents/itl/iad/mig/smartspace/Technical_Documentation_Mark_III.pdf (Last viewed May 2019).

J. Acoust. Soc. Am. **146** (6), December 2019

Madhu *et al.* 4911

Pan, G., Li, Y., and Chen, L. (**2019**). "Impact analysis and optimization of material parameters of insulator on brake squeal," IEEE Access **7**, 15861–15867.

Paulraj, A., and Kailath, T. (**1986**). "Eigenstructure methods for direction of arrival estimation in the presence of unknown noise fields," IEEE Trans. Acoust. Speech Signal Process. **34**(1), 13–20.

Philippen, B., Sottek, R., and Spiertz, M. (**2010**). "Blinde Quellentrennung von Fahrzeuginnengeräuschen" ("Blind source separation of in-car noises"), in *Proceedings of the Annual Meeting of the German Acoustical Society (DAGA)*, Berlin, pp. 111–112.

Renevey, P., and Drygajlo, A. (**2001**). "Entropy based voice activity detection in very noisy conditions," in *Proceedings of the European Conference on Speech Communication and Technology (EUROSPEECH)*, September 3–7, Aalborg, Denmark.

Roth, P. R. (**1971**). "Effective measurements using digital signal analysis," IEEE Spectrum **8**, 62–70.

Scheuing, J., and Yang, B. (**2006**). "Disambiguation of TDOA estimates in multi-path multi-source environments (DATEMM)," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 14–19, Toulouse, France, pp. 837–840.

Schmidt, R. O. (**1981**). "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. thesis, Stanford University, Stanford, CA.

Shannon, C., and Weaver, W. (**1949**). *The Mathematical Theory of Communication* (University of Illinois Press, Champaign, IL).

Slepian, D. (**1978**). "Prolate spheroidal wave functions, fourier analysis and uncertainty—V: The discrete case," Bell Syst. Tech. J. **57**, 1371–1430.

Tadaion, A. A., Derakhtian, M., Gazor, S., and Aref, M. R. (**2007**). "A fast multiple-source detection and localization array signal processing algorithm using the spatial filtering and ML approach," IEEE Trans. Signal Process. **55**(5), 1815–1827.

Talantzis, F., Constantinides, A. G., and Polymenakos, L. C. (**2005**). "Estimation of direction of arrival using information theory," IEEE Signal Process. Lett. **12**(8), 561–564.

Thiergart, O., Huang, W., and Habets, E. A. P. (**2016**). "A low complexity weighted least squares narrowband DOA estimator for arbitrary array geometries," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 20–25, Shanghai, China, pp. 340–344.

van Trees, H. L. (**2002**). *Detection, Estimation and Modulation Theory, Part IV* (John Wiley and Sons, New York).

Ward, D. B., Kennedy, R. A., and Williamson, R. C. (**2001**). "Constant directivity beamforming," in *Microphone Arrays: Signal Processing Techniques and Applications*, edited by M. Brandstein and D. Ward (Springer-Verlag, New York), pp. 3–17.

Wax, M. (**1992**). "Detection and localization of multiple sources in noise with unknown covariance," IEEE Trans. Acoust. Speech Signal Process. **40**(1), 245–249.

Wax, M., and Kailath, T. (**1984**). "Spatio-temporal spectral analysis by eigenstructure methods," IEEE Trans. Acoust. Speech Signal Process. **32**(4), 817–827.

Yoon, Y., Kaplan, L. M., and McClellan, J. H. (**2006**). "TOPS: New DOA estimator for wideband signals," IEEE Trans. Signal Process. **54**(6), 1977–1989.