

AUGUST 01 2011

Method for measuring violin sound radiation based on bowed glissandi and its application to sound synthesis

Alfonso Pérez Carrillo; Jordi Bonada; Jukka Pätynen; Vesa Välimäki



J. Acoust. Soc. Am. 130, 1020–1029 (2011)

<https://doi.org/10.1121/1.3605291>



Articles You May Be Interested In

Non-impulsive signal deconvolution for computation of violin sound radiation patterns and applications in sound synthesis

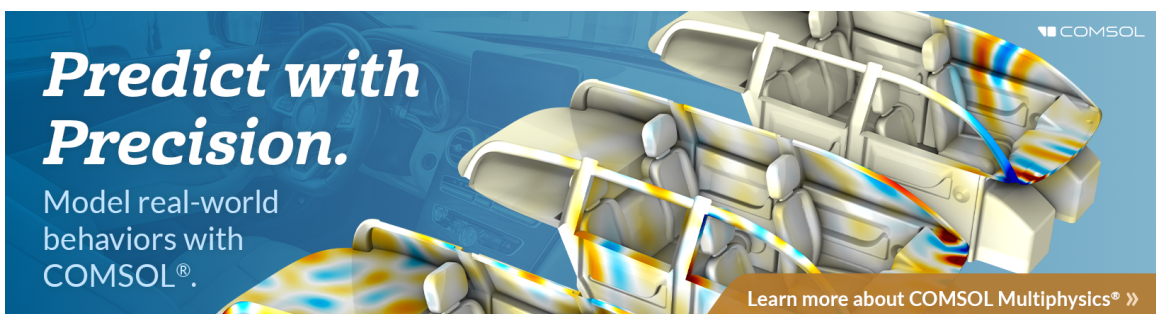
J. Acoust. Soc. Am. (September 2015)

Perception of complex-tone pairs mistuned from unison: Waveform relations determines whether pitch slides or iterated complex auditory patterns are heard

J. Acoust. Soc. Am. (August 2005)

A measurement study on voice instabilities during the register transition

J. Acoust. Soc. Am. (October 2016)



Predict with Precision.
Model real-world behaviors with COMSOL®.

COMSOL

Learn more about COMSOL Multiphysics® »

Method for measuring violin sound radiation based on bowed glissandi and its application to sound synthesis

Alfonso Pérez Carrillo^{a)} and Jordi Bonada

Music Technology Group Universitat Pompeu Fabra, Roc Boronat 138, 08018 Barcelona, Spain

Jukka Pätynen

Department of Media Technology, School of Science, Aalto University, P.O. Box 15400, FI-00076 Aalto, Espoo, Finland

Vesa Välimäki

Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, P.O. Box 13000, FI-00076 Aalto, Espoo, Finland

(Received 12 October 2009; revised 15 May 2011; accepted 19 May 2011)

This work presents a method for measuring and computing violin-body directional frequency responses, which are used for violin sound synthesis. The approach is based on a frame-weighted deconvolution of excitation and response signals. The excitation, consisting of bowed glissandi, is measured with piezoelectric transducers built into the bridge. Radiation responses are recorded in an anechoic chamber with multiple microphones placed at different angles around the violin. The proposed deconvolution algorithm computes impulse responses that, when convolved with any source signal (captured with the same transducer), produce a highly realistic violin sound very similar to that of a microphone recording. The use of motion sensors allows for tracking violin movements. Combining this information with the directional responses and using a dynamic convolution algorithm, helps to improve the listening experience by incorporating the violinist motion effect in stereo. © 2011 Acoustical Society of America. [DOI: 10.1121/1.3605291]

PACS number(s): 43.75.De, 43.75.Zz, 43.75.Yy [NHF]

Pages: 1020–1029

I. INTRODUCTION

In a simplified model of violin sound production and radiation, we can consider two main vibrating elements: the bowed-string interaction, which is the nonlinear component, and a system comprised of all the elements in the signal path from bridge vibration to sound perception, which may be considered as nearly linear (Cremer, 1984). This linear part is often referred to as the violin body, but it is actually a coupled system whose main components are the bridge, the violin body and even the room. Understanding the properties of the violin body is of great interest as it largely influences the sound quality of the instrument. It is usually modeled as a linear filter (Farina *et al.*, 1995, 1998; Bissinger, 2006) by measuring its response to a known excitation and computing the body frequency response (BFR) or body impulse response (BIR). There exist many different methods to obtain BFRs, which are adapted to the particularities of the research. There can be found works aimed at the comprehension of specific physical phenomena (Jansson *et al.*, 1970; Cremer, 1984; Marshall, 1985; Jansson *et al.*, 1994; Wang, 2000) and perceptual features related to the properties of the violin body (Woodhouse, 2005; Fritz *et al.*, 2006, 2007). Some studies report progress in violin making (Schleske, 2002). Finally, there can be found works making use of BFRs for sound synthesis (Smith, 1983; Cook and Trueman, 1998; Vos *et al.*, 2003; Waxman, 2005). However, their

main objective has not been highly realistic synthesis, which is the goal of this work.

The aim of this study is to develop a general technique to obtain BFRs (and their respective BIRs) between the signal captured with a pickup on the violin and the radiated sound pressure in different directions relative to the body. The pickup is built into the bridge and we consider both the bridge and the pickup to be part of the body. The intention of our approach is not physical modeling, so the precise point in the signal path at which the pickup is mounted and the correlation between the pickup signal and some physical quantity are irrelevant. The main requisite for the obtained BIRs is that they must produce a highly realistic sound when convolved with a source signal [recorded with the pickup or synthesized with a signal model (Perez, 2009)]. The proposed methodology is based on bowed glissandi that are recorded with the pickup (excitation) and several microphones (responses). The frequency responses are obtained by computing the transfer functions between the pickup and each microphone as a weighted average of the deconvolution of short temporal frames. From the measurement and analysis described above, the estimated BFRs are incorporated into an enriched sound synthesis of violin performance. The use of BFRs allows us to imitate the sound of different violins as in the work by Fritz *et al.* (2006), simulate stereo listening as reported in Waxman (2005) and simulate different positions of the listener. Furthermore, provided with a motion sensor that tracks the position of the violin, we can simulate the effect of a moving performer. This is achieved by applying the convolution at each instant with the respective BIR.

^{a)}Author to whom correspondence should be addressed. Electronic mail: alfonso.perez@upf.edu

II. RELATED WORK

Many different methods are used to obtain the violin BFR, usually based on measuring the body's response to a known excitation. They can be differentiated by the place where the excitation is applied and type of excitation signal, by the place where the response is measured and the type of response, and by the algorithm to compute the transfer function from excitation and response signals, which typically is based on signal deconvolution (Smith, 1997) or system identification (Smith, 1983).

The excitation signal may be impulsive or continuous. Impulsive excitations can be performed by hitting the violin bridge with a force hammer (Woodhouse, 2005). The impulse may be applied directly to the strings in order to obtain a different BFR for each string (Cook and Trueman, 1998). Non-impulsive excitations, e.g., sine sweep or MLS (Farina *et al.*, 1998), are usually less affected by noise and measurement errors and may be applied by means of a shaker or wire force transducer (Farina *et al.*, 1998) or following the "indirect" method with a loudspeaker (Weinreich and Arnold, 1980; Farina *et al.*, 1995; Vos *et al.*, 2003). Other possible methods are finger tapping, traditionally used by violin makers, or playing the instrument as in Karjalainen *et al.* (2000), where a guitar is excited by playing musical excerpts.

Violin body transfer functions can be mechanical or acoustical depending on whether the response is measured as mechanical vibration or sound pressure. Mechanical vibration is usually measured as velocity or acceleration at the bridge with an optical laser (Woodhouse, 2005), with an accelerometer (Morset, 2001), or with a magnet inside a coil (Jansson, 2002). Acoustical responses are measured with a microphone, ideally in an anechoic chamber in order to avoid the room reverberation.

In this work, sound radiation is measured at different angles around the violin, making it possible to obtain radiation patterns (Weinreich and Arnold, 1980; Williams *et al.*, 1980; Cook and Trueman, 1998; Otčenášek and Syrový, 1999; Wang and Burroughs, 2001; Bissinger, 2008a). Weinreich (1997) and Bissinger (2008b) have extensively studied the directional radiation from violins. In general, previous studies concentrate on a limited range of low frequencies but for our synthesis purposes we need to obtain the response over the full audio frequency range.

Violin sound synthesis based on a source plus filter approach is often used in the field of violin acoustics, but the final objective has not been highly realistic synthesis, as is the case in our work. For example, Otčenášek and Syrový (1999) report a perceptual analysis about directivity of the violin timbre, and Fritz *et al.* (2006) carry out psychoacoustic experiments with *virtual* violins. Holm and Välimäki (2000) investigate modifications of the resonances of violin body modes and Smith (1983) applies filter design techniques to model violin body resonances.

III. METHODOLOGY

The methodology for obtaining the violin BFR is based on a similar procedure as reported by Karjalainen *et al.*

(2000), who present a method to compute guitar acoustic transfer functions. They excite the guitar by playing tones and measure the excitation with an under-saddle pickup and the response with a microphone, and finally compute the BFR by signal deconvolution. Our approach is based on playing glissandi that were recorded with an instrument pickup (excitation) and a microphone (response), and we estimate the BFR with an algorithm based on signal deconvolution.

A. Excitation and response measurement

Several typical excitation methods were tried and abandoned in favor of playing musical excerpts with the bow. The main reason is that common methods for obtaining BFRs are based on applying a known excitation signal (e.g., an impulse, a pseudo-random noise sequence, or a sine sweep), but in our case, the excitation signal is not the applied signal, but the signal captured by the pickup. Furthermore, the commonly used methods are not adapted to our specific needs. For instance, by striking the instrument with a force hammer, we obtained a non-noisy BFR up to around 10 kHz and the noise of the hammer hitting the bridge becomes a problem as it is captured by the microphone. The use of a shaker adds an extra mass to the bridge, and we could only obtain an acceptable response at low frequencies. Finally, it was decided to excite the violin by bowing as it does not interfere when measuring the response with a microphone and it is the natural way of exciting the violin. Moreover, the measurement conducted this way includes the influence of the performer, which is desirable for our synthesis purposes. Playing glissandi was preferred rather than performing other musical exercises, as it is a simple method to sweep and homogeneously excite all frequencies in a wide frequency range.

Glissandi were performed on the G-string, beginning at the open string and with a range slightly wider than one octave. Using the G-string we obtain the lowest playable pitch of 196 Hz on the violin with the tone G3. A glissando over the lowest playable octave is sufficient for exciting the entire frequency range of interest, and no significant difference in sound quality was perceived when using BFRs obtained from glissandi with higher pitch extensions. It is possible to estimate individual BFRs for each string. However, this idea was abandoned for simplification and because the result of the convolutions was perceptually found to be nearly identical according to informal listening. Glissandi lasted for approximately 50 s as they must be played slowly in order to be nearly stationary during each temporal frame (see next subsection). They were performed with an almost constant rate of increasing pitch (semitones/s) that would approximately correspond to a logarithmic frequency sweep. Approximately fifteen bow strokes were necessary to complete an entire glissando.

The excitation was measured with a pickup. Several sensors and commercial pickups were tested and the selected device¹ (a piezoelectric pickup integrated with the bridge) was chosen on basis of the quality of the corresponding convolution with a source signal. The response was measured with several Røde-NT1A microphones, and a compensation filter was applied in order to achieve an ideal flat response in

the same manner as in Pätynen *et al.* (2008). We evaluated the spectral richness of the excitation glissando from histograms of the energy contribution to each spectral bin. In Fig. 1 we show the resulting histogram for a glissando recorded with the bridge pickup and compare it to the spectrum of one of the first frames of the same glissando. The spectrum envelope falls smoothly 70 dB until 10 kHz. A clear partial structure is present in the spectrum of the single frame up to that frequency. Above 10 kHz the spectrum level is approximately flat.

Both distance and orientation of the microphones relative to the violin affect the properties of the recorded sound, and it is desirable to obtain a consistent estimation after time-averaging individual frame estimations. For this reason, we built a mechanical structure to keep the violin fixed, preventing the performer from moving the instrument while playing the glissando.

B. Energy weighted multiple frame deconvolution

Typical ways of obtaining violin body frequency responses are by means of system identification techniques (Smith, 1983) or by signal deconvolution (Farina *et al.*, 1998; Karjalainen *et al.*, 2000). Applying an *exact* deconvolution by division in frequency domain to signals, such as the proposed glissando, would result in a very noisy frequency response. There are a few reasons for this. First, the experimental measurement has limited frequency resolution, and further, the presence of noise and the use of a simplified model ignoring potential non-linearities and time-variance lead to inaccuracies. Hence, we compute the deconvolution on a frame-by-frame basis followed by a global averaging across all frames of the glissando in which each frame-deconvolution is weighted according to the energy of the excitation signal in that frame.

The signal recorded with the microphone has a delay of approximately 6 ms, corresponding to the distance of 2 m from the violin. First, this delay is removed to align excitation $x(t)$ and response $y(t)$ signals. Then, both signals are windowed and expressed in the spectral domain, obtaining two frame streams that are aligned in time $[X(f,k)$ and $Y(f,k)$, where f is the frame and k is the spectral bin]. At this point, we apply a frame-by-frame deconvolution. The magnitude is obtained as

$$|H_{\text{BFR}}(f, k)| = \frac{|Y(f, k)|}{\max(|X(f, k)|, b)}, \quad (1)$$

where the value 10^{-5} was chosen for the threshold b . The phase is obtained as

$$\phi[H_{\text{BFR}}(f, k)] = \phi[Y(f, k)] - \phi[X(f, k)] + 2\pi m, \quad (2)$$

where m is the integer that wraps the phase difference in the interval $[-\pi, \pi]$. Now, having estimated the BFR for individual frames, separate BFRs are averaged in order to obtain an accurate estimation of the overall BFR.

For the magnitude estimation, we use the typical weighted average expressed as

$$|H_{\text{BFR}}(k)| = \frac{\sum_{1 \leq f \leq N} w_f(k) \times 20 \log_{10} |H_{\text{BFR}}(f, k)|}{\sum_{1 \leq f \leq N} w_f(k)} [\text{dB}], \quad (3)$$

where f is the frame number, N is the number of frames, k is the spectral bin being estimated, and $w_f(k)$ is the weight being applied, which corresponds to the input signal energy at frame f and bin k .

Regarding the phase, due to its cyclic behavior, carrying out a classical weighted average would not provide good estimations. As a first attempt, we explored a method based on constructing a histogram of the phase values estimated for each spectral bin, weighted by their corresponding energy. However, the resulting BIRs were not causal, so finally we computed the minimum phase BIRs from the estimated BFR magnitudes by using the cepstrum and converting anti-causal exponentials to causal exponentials (Oppenheim and Schaffer, 1975).

Glissandi were recorded at a sample rate of $f_s = 44100$ Hz and spectrally analyzed with the following parameters. A Blackman-Harris window with length of 8192 samples (186 ms) is used with a window overlap of 50% and a zero-padding factor of 8 (thus a FFT size of 65536). With these analysis parameters we achieve a bin resolution in the computed BFR of approximately 0.67 Hz. The effective length of the resulting BIRs corresponds to half window length, that is, 93 ms.

The proposed deconvolution algorithm assumes that the glissando is quasi-stationary within the analysis window in

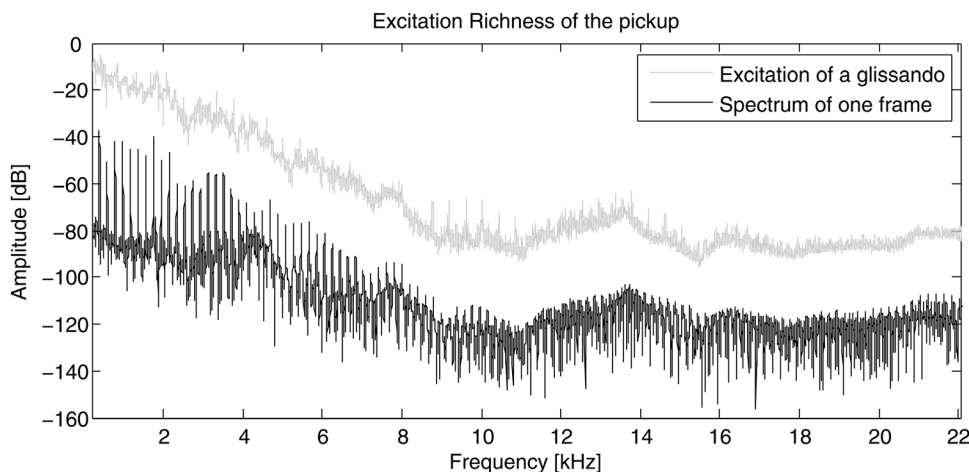


FIG. 1. Observation of the spectral content of the input excitation signal from the bridge pickup used for deconvolution. Histogram maxima for all spectral bins across a glissando (gray) are compared with the spectrum of one of the first frames of the same glissando (black).

order to assure that the energy of the short-time spectrum of an arbitrary frame is mostly concentrated around harmonic frequencies. Therefore, the use of a relatively short analysis window (around 186 ms) implies that a slow glissando must be used as the input signal. In order to illustrate the effect of this energy distribution, let us consider an ideal case where the input signal $x(t)$ is a complex exponential at frequency f_0 , amplitude A , and initial phase φ_0

$$x(t) = A_0 e^{j(2\pi f_0 t + \varphi_0)}. \quad (4)$$

When this signal is filtered, the output signal obtained $y(t)$ is equal to the input multiplied by a phasor determined by the filter frequency response,

$$y(t) = A_0 e^{j(2\pi f_0 t + \varphi_0)} H(f_0), \quad (5)$$

where H is the frequency response of the filter.

In practice we must consider discrete signals, obtained by sampling at a fixed rate of f_s samples per second. The discrete Fourier transform (DFT) of a windowed complex exponential is given by

$$\begin{aligned} X_w(k) &= \sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi kn/N} \\ &= \sum_{n=0}^{N-1} w(n)Ae^{j(2\pi k_0 n/N + \varphi_0)} e^{-j2\pi kn/N} \\ &= Ae^{j\varphi_0} \sum_{n=0}^{N-1} w(n)e^{-j2\pi(k-k_0)n/N} \\ &= Ae^{j\varphi_0} W(k-k_0), \end{aligned} \quad (6)$$

where n is the sample index, k the spectral bin, N the length of the DFT, w the analysis window, W its transform, and $k_0 = (f_0/f_s)N$. The DFT is equal to the phasor of the exponential multiplied by the transform of the window shifted to the frequency of the exponential. The estimated filter \hat{H} is computed by dividing the Fourier transforms of the windowed output and input discrete signals,

$$\begin{aligned} \hat{H}(k) &= \frac{Y(k)}{X(k)} = \frac{\sum_{n=0}^{N-1} w(n)y(n)e^{-j2\pi kn/N}}{\sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi kn/N}} \\ &= \frac{A_0 e^{j\varphi_0} W(k-k_0)H(k_0)}{A_0 e^{j\varphi_0} W(k-k_0)} = H(k_0), \end{aligned} \quad (7)$$

that is, the estimated filter is constant for all spectral bins, and equal to the value of the actual filter at the exponential frequency. If we consider as input two complex exponentials of frequencies k_0 and k_1 cycles per sample, the estimated filter becomes

$$\begin{aligned} \hat{H}(k) &= \frac{Y(k)}{X(k)} = \frac{\sum_{n=0}^{N-1} w(n)y(n)e^{-j2\pi kn/N}}{\sum_{n=0}^{N-1} w(n)x(n)e^{-j2\pi kn/N}} \\ &= \frac{A_0 e^{j\varphi_0} W(k-k_0)H(k_0) + A_1 e^{j\varphi_1} W(k-k_1)H(k_1)}{A_0 e^{j\varphi_0} W(k-k_0) + A_1 e^{j\varphi_1} W(k-k_1)}, \end{aligned} \quad (8)$$

which is a weighted sum of the actual filter values at the exponential frequencies, where the weighting depends on the transform of the analysis window. For typical windows the energy of the transform is concentrated around its main lobe. Consequently, the estimated filter is approximately constant around each exponential frequency.

When we combine several frame estimations as discussed above, each harmonic contributes to the averaged response with a weight, which is determined mostly by its energy and the transform of the analysis window. Hence, the ratio between output and input signals can be considered as a sampling of the target filter at harmonic frequencies. This is shown in Fig. 2.

If the analysis window is longer or the glissandi are performed faster it may happen that the excitation is no more quasi-stationary within the analysis window. This would have the effect of widening the harmonic peaks especially at mid and high frequencies (the harmonic frequency shift is larger at high frequencies than at low frequencies), negatively affecting the estimation of the filter. This is so due to the fact that we are performing a short-time analysis with a window that does not include the whole excitation and response signals. Therefore, it is recommended to use slow glissandos. Besides, for a good filter estimation, it is necessary that the sweeping harmonics along the whole glissando excite all the spectral bins in order to avoid zeros in the excitation spectrum.

In Fig. 3 we show the obtained BFR magnitude at the same microphone and violin position for four different glissando recordings to show the repeatability of the proposed method. BFRs are seen to be very similar, although there can be found differences among them that seem to vary along frequencies. The reason is that the responses are affected by the exact position of the violinist when playing, especially his head.

IV. MEASUREMENTS

The objective of the measurements is to obtain sound radiation levels around a violin at all frequencies covered in normal playing by bowing. This is achieved by computing the transfer function between the bridge pickup and a set of microphones placed at different angles. BFRs can be convolved subsequently with a source signal in order to simulate

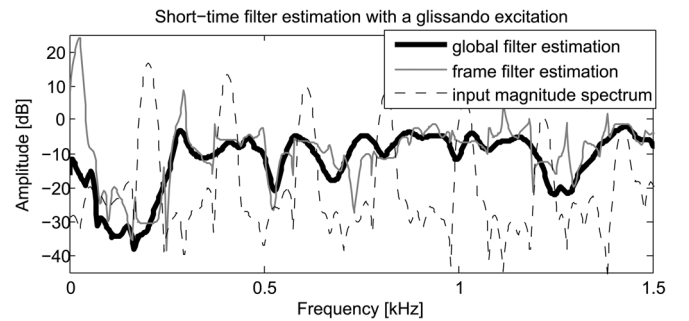


FIG. 2. Short-time filter estimation (bold) with a glissando excitation. The filter estimation at each frame of the glissando [$H_{\text{BFR}}(f, k)$, thin curve] has almost constant values around the center frequency of each harmonic of the input magnitude spectrum of that frame [$|X(f, k)|$, dotted curve].

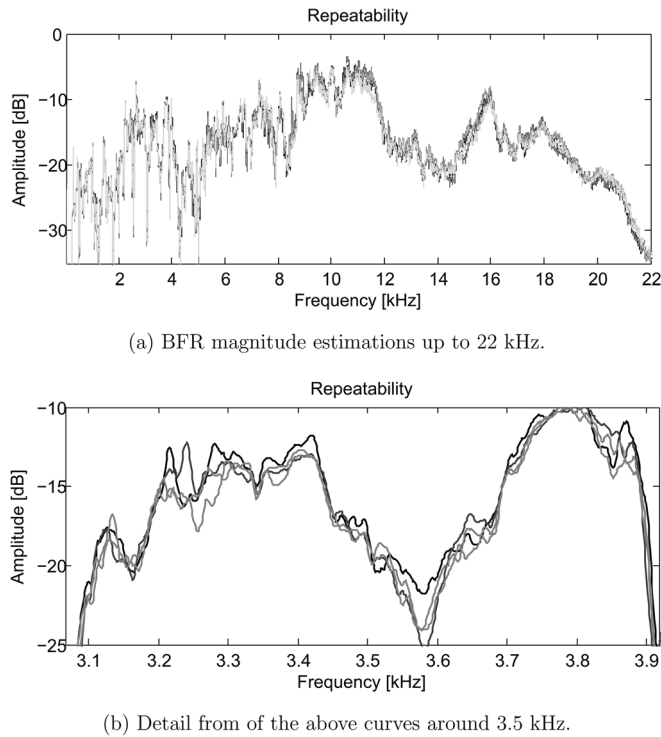


FIG. 3. Repeated computation of the BFR magnitude of four different glissando recordings at the same angle and distance violin-microphone to show the repeatability of the estimations. BFRs are seen to be very similar and differences may be caused by the exact position of the violinist when playing, especially his head.

the position of the listener with respect to the violinist. The BFR at each angle is estimated as described in the previous section. The measurements were performed using a violin of professional quality.

A. Setup and calibration

The measurements were conducted in an anechoic chamber at Aalto University with anechoic conditions down to about 125 Hz. Setup and microphone calibration were performed following the same basic concept as in the work by Pätynen *et al.* (2008). The measurement room was equipped with 21 Røde microphones placed around the center of the room in a spherical-like arrangement (two pentagonal anti-prisms) at distances from the center between 1.7 and 2.4 m. The microphones were grouped in four horizontal planes at nominally 10 and 50 deg above and below the equatorial plane, with an additional microphone close to the top. In each horizontal plane there were five microphones evenly spaced (around 72 deg between consecutive microphones). All angles and distances relative to the center of the room were accurately measured with laser equipment.

Due to the different positions of the microphones in relation to the absorbing wedges on the walls, varying distances from the room center, and the inherent differences in frequency response between microphones, the same procedure as in Pätynen *et al.* (2008) was performed in order to equalize the responses of the recording microphones.

We built a mechanical structure to prevent the performer from moving the violin while playing the glissandi.

The structure is thin to minimize its influence on the measurements and it can hold the violin with two standard shoulder rests (Fig. 4). The support of the violin may be assumed to be similar to normal playing conditions when the performer is standing.

B. Coordinate systems

The violin was placed at the center of the room by defining 3D coordinate systems for the violin and room, respectively, and matching the origins of the two systems with the help of an *OptiTrack* system² with six cameras.

The origin of the room system was set at the center of the room and was determined with the help of a laser pointer and marks on the walls at the same height, separated by 90 deg. The room x - y plane was the equatorial plane of the microphone array, which is parallel to the floor (measured with levels).

The origin of the violin coordinate system (see Fig. 4) was defined as the midpoint between the projections of the E and G string-to-bridge terminations on the top plate plane. The violin x axis was in the direction of the neck with the x - y plane being parallel to the top plate.

The violin was mounted on the supporting structure and fluorescent markers were attached to it (Fig. 4). With the *OptiTrack* visualization tool we make violin and room coordinates coincide, making it possible to measure the angle differences between the two coordinate systems.

The supporting structure can be easily rotated to fix the violin at a specific orientation (see Fig. 4), determined by the violin azimuth (az_v) and inclination (inc_v) with respect to the

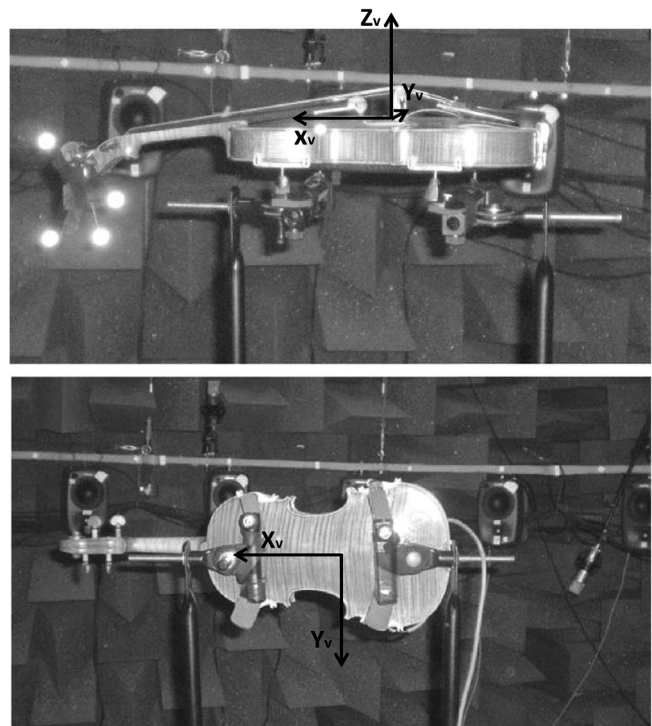


FIG. 4. Violin with markers mounted in the mechanical structure. Violin inclination $inc_v = 0$ deg (top) and $inc_v = 90$ deg (bottom). Violin coordinate system is shown with black arrows. Violin elevation el_v was kept at 0 deg in all measurements.

room system (Euler angles). The azimuth is changed through rotation of the violin around the z -axis (by rotating the supporting structure around the vertical axis of the stand), and the inclination through rotation around the violin x axis (by rotating around the pivoting axis of the shoulder rest). Note that the shoulder rest axis is parallel to, but not coinciding with the violin x axis, so each time the inclination is changed, violin and room systems need to be realigned. In addition to the *OptiTrack* system, the violin azimuth is measured with a protractor, placed at the base of the supporting structure. Violin elevation (el_v) was kept constant in the measurements at 0 deg, meaning that the violin z -axis always coincided with the equatorial plane of the microphone array. In Sec. V B the sampled microphone (or listener) positions are expressed in spherical coordinates with respect to the violin, which are determined by the listener azimuth (az_l) and elevation (el_l), see Fig. 8. Note that moving the violin by rotating the structure, is equivalent to moving the listener to the opposite direction.

C. Directional radiation

Radiation is understood as the quantity of radiated sound energy, which varies spatially around the violin. Directivity is usually defined as the *radiation ratio* among different directions at a certain frequency (Weinreich, 1997). As a result of the measurements, we obtained several BFRs around the violin. Although the estimated BFRs are not directly related to a physical property of the violin (because of the pickup), the ratio between BFRs shows the radiation ratio at different directions.

The measurements were conducted as follows. The violin was mounted in the structure and set to a certain inclination (inc_v). For that inclination, the 360 deg of azimuth (az_v) were sampled in steps of 5 deg. As each horizontal plane has five microphones, it was sufficient to record 15 glissandi at steps of 5 deg going from one microphone to the next in the same horizontal plane. The same process was repeated for inclinations at 0, 30, 60, and 90 deg in order to obtain the sampled space represented in Fig. 5, in all including 1260 sampled positions (each cross indicates a sampled position). Violin elevation was kept constant at 0 deg, as mentioned. However, by measuring the violin with different azimuth and inclination angles, also changes in elevation can be derived from the performed measurements.

Figure 6 shows the radiation ratio between two randomly selected directions (with the same el_l and separated by 72 deg az_l). Although the ratio is different for each pair of directions, the plot is representative for the general behavior. Below 600 Hz, radiation is nearly identical in both directions (the difference is around 1 dB). From 600 to 1000 Hz the radiation ratio is small although some peaks are present. The highest differences appear in the range 1000–4500 Hz, reaching in some cases 20 dB.

In Fig. 7 we visualize 2D polar directivity curves in dB at specific frequency ranges measured with five microphones located in the same horizontal plane (10 deg above the equatorial plane of the microphone array). The center of each polar graph (0 dB) represents the minimum radiated value at

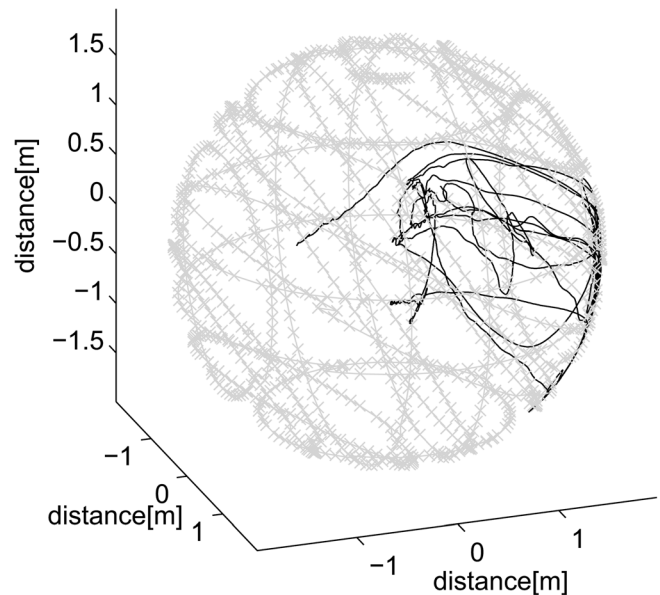


FIG. 5. Spherical position of sampled BFRs around the violin, in all 1260 positions (gray crosses) and listener trajectory (black line) with respect to the stationary violin, recorded during a real performance.

that frequency. Curves close to the center (i.e., a short radius) mean small differences in energy radiation among different angles, while a long radius indicates a high radiation ratio.

The graphs in Fig. 7(a) (violin inclination 0 deg) show the radiation in a plane parallel to and 35 cm above the x - y plane and those in Fig. 7(b) (inclination 90 deg) in a plane parallel to and 35 cm to the left of the x - z plane (bass bar side). The directivity patterns are seen to change greatly,

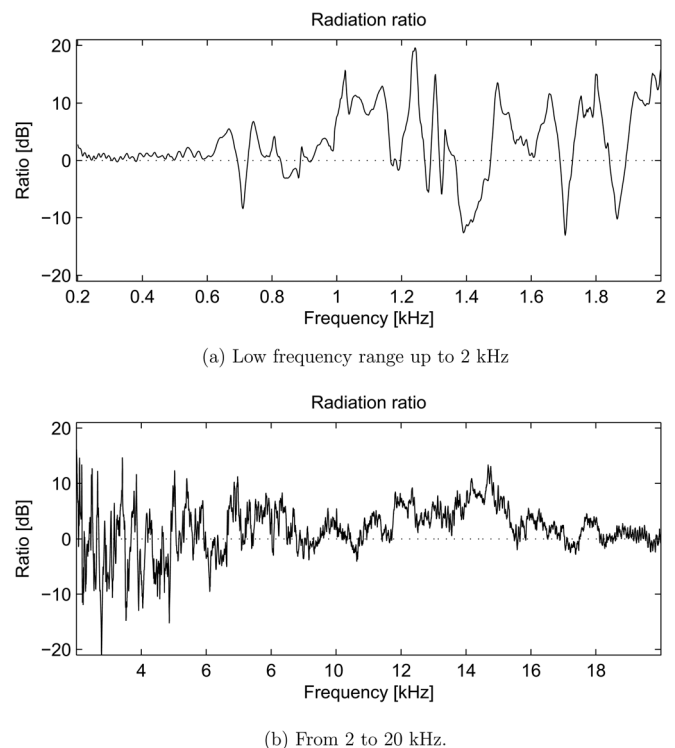
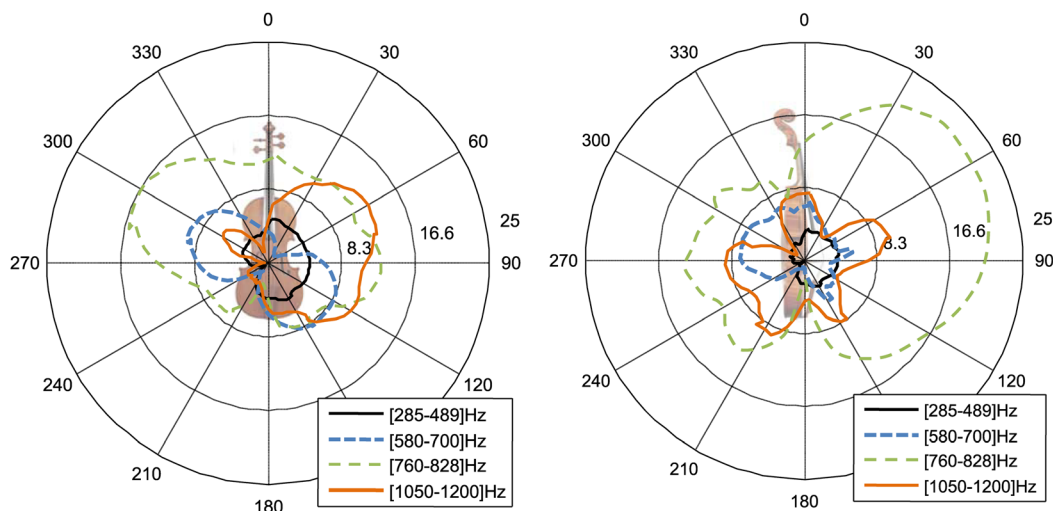


FIG. 6. Radiation ratio (dB) between two randomly selected directions. In this case the directions differ 72 deg in az_l and have the same el_l .



(a) Directivity patterns in a plane parallel to and 35 cm above the x - y plane. (b) Directivity patterns in a plane parallel to and 35 cm to the left of the x - z plane (bass bar side).

FIG. 7. (Color online) Violin directivity patterns.

even between very close frequencies. A video showing the evolution of the patterns along the frequency can be found in a dedicated web page.³

In both planes in Fig. 7 we can observe how at low frequencies (196–489 Hz) directivity is very low. The radiation is almost omnidirectional except at the position of the player who introduces a shadowing effect. In the x - y plane in the range 580–700 Hz, shapes seem to be symmetrical with respect to an axis at -60 deg, forming a clear pattern-of-eight which indicates a dipole source. Again, lowest radiation values are found at the position of the violinist. In the range 760–828 Hz most of the energy is radiated in the entire frontal direction of the violin, and in the range 1050–1200 Hz to the right side. For the x - z plane in the range 580–700 Hz, most of the energy is radiated from the back, whereas in the range 760–828 Hz it is rather symmetrical with respect to the violin x axis (neck direction), but with a stronger radiation from the top plate. The radiation in both directions of the x axis is low. At higher frequencies 1050–1200 Hz we observe a five-arm star shape.

In general, the observed behavior is consistent with previous studies (Weinreich, 1997). The main difference is that the performer is included in our measurements. It can clearly be observed that the lowest radiation values appear in the direction he is standing.

D. Discussion

The brief description of the radiation patterns above shows that the observations are in agreement with previous studies (Weinreich, 1997). A more comprehensive analysis of the patterns will be presented elsewhere, as in this presentation we want to focus on their application for sound synthesis. Furthermore, the observed results were obtained with only one violin of professional quality. By analyzing a set of violins of different quality we may find specific features that

distinguish fine violins from others and, at the same time, we may find properties common to all violins, which would allow for defining a set of *generic* violin BFRs.

V. APPLICATION TO SOUND SYNTHESIS

Given a source signal, either recorded with the bridge-pickup or synthesized (Perez, 2009), we can obtain a realistic sound close to that captured by a microphone, by convolving it with the estimated body impulse responses (BIRs). Convolution is performed via the frequency domain following the procedure reported in Arfib *et al.* (2002). A rectangular window is used to select an input signal frame of the size of the BIR, which is zero padded and multiplied in the frequency domain with a zero padded BFR. The result, which is twice the length of the window, is transformed back to the time domain and overlap-added with previous data. The window is shifted by the length of the BIR and the same procedure is repeated.

A. Simulation of stereo and performer motion

As we computed several BFRs around the violin we can recreate the sound perceived by a listener at different directions and provide a better listening experience by simulating effects such as stereo or the movement of the player. As a simplification we consider the listener to be seated still at a distance of 2 m (the same distance as from the microphones to the violin during the measurements), and we assume that the ear-to-ear angle difference as viewed from the violin can be approximated as five degrees at that distance. In this manner, stereo can be simulated directly by using two consecutive BIRs in any horizontal plane, one for each ear.

The movement of the performer is another effect that is necessary to be included in a realistic simulation. Ancillary gestures are constantly changing the position and orientation of the violin, so the listener is perceiving the sound radiated

at continuously changing directions. To simulate this motion we need to know the position of the listener relative to the violin and apply the convolution with the corresponding BIR at each instant. This effect is more clearly perceived in stereo, so the same procedure must be performed for both ears. In order to demonstrate this effect, we made use of a multimodal database of real violin performances that contains recordings of the pickup signal together with the position and orientation of the violin. Any recorded pickup signal may be then convolved with the BFRs corresponding to the coordinates of the violin. The method is briefly explained in the following subsection and some sound samples are found in a dedicated web page.³

B. Listener coordinates

The position of the listener with respect to the violin is defined by the listener azimuth (az_l) and elevation (el_l) for each ear. Violin coordinates during performance were acquired with a Polhemus⁴ 3D-motion tracker, consisting of an electro-magnetic field source and a small sensor attached to the violin. The system tracks the position and orientation of the sensor with respect to the source with a frame rate of 240 samples/s. Therefore, the problem consists of converting the recorded sensor coordinates to listener coordinates with respect to the violin.

Let us consider three coordinate systems (Fig. 8), the *source* given by the axis vector $[X, Y, Z]$, the *sensor* (S) with axes $[X_s, Y_s, Z_s]$ and the *violin* (V) with axes $[X_v, Y_v, Z_v]$. The sensor was attached to the violin, so their coordinate systems have a fixed position and orientation with respect to each other. Similarly we consider the position of the listener relative to the source to be constant. Each coordinate system is represented as a tuple with position and rotation matrices.

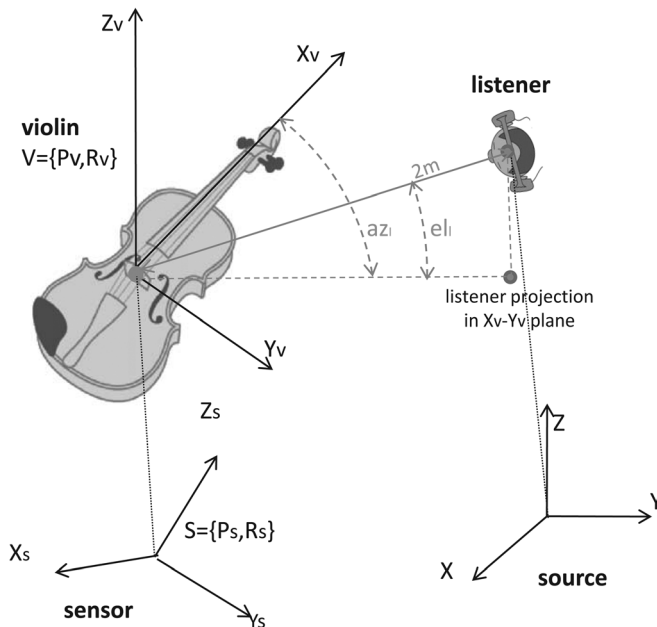


FIG. 8. Violin coordinate system $[X_v, Y_v, Z_v]$ and angles used to obtain microphone (listener) position relative to the violin coordinate system: azimuth (az_l) and elevation (el_l).

Coordinates of a point in any system can be expressed in a second system by knowing the tuple of the second system with respect to the first.

Using standard techniques, updated violin coordinates are computed relative to the source system for every new position and orientation of the sensor [see [Maestre et al. \(2007\)](#) and [Perez \(2009\)](#) for details]. The listener position is then translated to the new violin coordinate system, and finally converted to spherical coordinates (az_l and el_l) with respect to the violin system. Figure 5 shows the trajectory of the listener relative to the violin during a performance, together with the position of all sampled BFRs.

C. BIR interpolation

Listener position coordinates may not coincide with any of the estimated BFRs, so we need to interpolate among the closest measured locations. Euclidean distance is computed to each sampled BFR in Cartesian coordinates. The n closest BFRs to the listener are selected ($n = 10$ was used in our tests). The interpolation is a weighted sum of the BIRs (in the time domain), where the weight of each BIR is inverse to the distance and the sum of all weights is unity.

D. Dynamic convolution

In order to simulate the effect of the violinist moving we need to apply a convolution with a different BIR at each instant (the BIR corresponding to each new listener position). If we follow the previously introduced convolution algorithm [Arfib et al. \(2002\)](#), the maximum rate at which a change of BIR can be applied is given by its duration (93 ms). This may result in significant sound changes between consecutive convolutions. If we want to simulate faster changes in the radiation characteristics we cannot use full-length BIRs. We propose to use a dynamic convolution algorithm that involves breaking up a long BIR into smaller chunks and cross-fading among the smaller BIRs in a similar way to [Foster et al. \(1991\)](#), [Gardner \(1995\)](#) and [Garcia \(2002\)](#).

BIRs are divided into n sub-BIRs, and each frame is convolved with all n -fragments of its corresponding BIR, and finally the sum of all convolutions is computed. When there is a BIR change, the current frame has to be convolved with both the previous and current BIRs and cross-faded. Crossfading is performed at each frame: If there is no change of BIR between two frames, the result is the same as if the crossfade was not applied. Let n be the number of sub-BIRs, f_i the input frame, $h_{i,j}$ the j th sub-IR of the BIR at frame f_i and $f_{cr}(in, out)$ the crossfade function. The value of the output i th frame f_i^{out} is

$$f_i^{out} = \sum_{k=0}^{n-1} f_{cr}(f_{i-k} * h_{i-k,k}, f_{i-k} * h_{i-k-1,k}). \quad (9)$$

The crossfade function is defined in Eq. (10). $a(t)$ is faded-in and $b(t)$ faded-out, l_{subIR} is the length of the sub-IR and t is time:

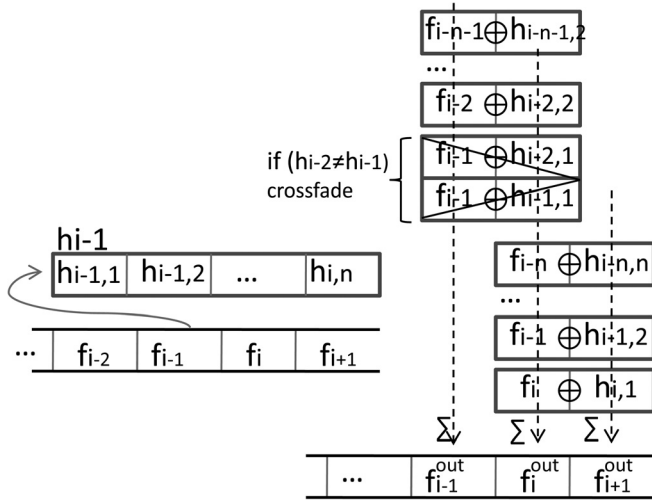


FIG. 9. Dynamic convolution scheme.

$$cf[a(t), b(t)] = \begin{cases} a2t + b(1 - 2t), & 0 < t \leq \frac{l_{subIR}}{2} \\ a, & \frac{l_{subIR}}{2} < t \leq l_{subIR}. \end{cases} \quad (10)$$

Figure 9 shows the process to compute the proposed dynamic convolution. Each output frame is twice the size of the input frame as a result of the convolution, thus they are overlap-added.

Some sound samples can be found in a dedicated web page.³ In these samples the number of sub-BIRs is 16 (2^4), and the length of the sub-BIRs (l_{subIR}) is the length of the whole BIR (2^{14} samples) divided by the number of sub-BIRs, that is, $l_{subIR} = 2^{10}$.

VI. CONCLUSION

In this paper we have introduced a new method to compute transfer functions between a violin bridge-pickup signal and radiated sound by playing glissandi, which are recorded simultaneously with a bridge-pickup (excitation) and a microphone in an anechoic chamber (response). The body impulse response is obtained by performing a frame-weighted deconvolution of the pickup and microphone recordings. Using multiple microphones, it is possible to simultaneously measure a set of body impulse responses in various directions around the violin. The impulse responses can be convolved with a source signal obtained with the bridge-pickup in order to obtain realistic violin sounds. Alternatively, the source signal can be generated using a string sound synthesizer. In this case, the method enables improvement of the sound quality of the synthesizer and the inclusion of directivity in sound synthesis. Acoustic radiation measurements conducted in this way are not restricted to a specific frequency range. Although the method is used here for the violin, this is a general procedure that can be applied to other musical instruments as well. The requirement is that there must be a way to measure an excitation signal together with the response signals.

The multiple impulse responses at different directions make it possible to enhance sound synthesis with spatializa-

tion effects. Performances can be recorded while keeping track of the position of the violin, thereby enabling the simulation of the moving effect of the performer. To demonstrate some of these possibilities, simulations of stereo listening and performer motion were presented. Additionally, we proposed an algorithm for dynamic convolution, in which each impulse response is divided into short frames, allowing for real-time cross-fading of impulse responses in a virtually continuous manner.

ACKNOWLEDGMENTS

This research received partial support by AGAUR (Catalan Government), by the Academy of Finland under Project 122815 and by Yamaha Corp.

¹Yamaha VNP1 bridge pickup system. <http://usa.yamaha.com/products/musical-instruments/strings/accessories/violin-pickups/vnp1/?mode=model> (date last viewed 10/23/10).

²Optitrack-optical motion capture systems. <http://www.naturalpoint.com/optitrack/> (date last viewed 10/23/10).

³Dedicated web site with extra information and sound samples. <http://www.dtic.upf.edu/~aperez/JASAradiation/> (date last updated 04/21/10).

⁴Polhemus Liberty System http://www.polhemus.com/?page=Motion_Liberty (date last viewed 10/23/10).

Arfib, D., Keiler, F., and Zölzer, U. (2002). *DAFX - Digital Audio Effects*, chapter 8.4.1: Time-frequency Filtering (Wiley, Chichester, England), pp. 263–264.

Bissinger, G. (2006). “The violin bridge as filter,” *J. Acoust. Soc. Am.* **120**, 482–491.

Bissinger, G. (2008a). “Structural acoustics model of the violin radiativity profile,” *J. Acoust. Soc. Am.* **124**, 4013–4023.

Bissinger, G. (2008b). “Structural acoustics of good and bad violins,” *J. Acoust. Soc. Am.* **124**, 1764–1773.

Cook, P. R., and Trueman, D. (1998). “A database of measured musical instrument body radiation impulse responses, and computer applications for exploring and utilizing the measured filter functions,” in *Proc. Int. Symp. Musical Acoust.* (Acoustical Society of America, Woodbury, NY), pp. 303–308.

Cremer, L. (1984). *Physics of the Violin* (The MIT Press, Cambridge, MA), pp. 201–382.

Farina, A., Langhoff, A., and Tronchin, L. (1995). “Subjective comparisons of ‘virtual’ violins obtained by convolution,” in *Proc. 2nd Int. Conf. on Acoustics and Musical Research*, Ferrara, Italy, pp. 36–41.

Farina, A., Langhoff, A., and Tronchin, L. (1998). “Acoustic characterisation of ‘virtual’ musical instruments: using MLS technique on ancient violins,” *J. New Music Res.* **27**, 359–379.

Foster, S., Wenzel, E., and Taylor, R. (1991). “Real time synthesis of complex acoustic environments,” in *Proc. IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, pp. 47–48.

Fritz, C., Cross, I., Moore, B. C. J., and Woodhouse, J. (2006). “Psychoacoustic experiments with virtual violins,” *J. Acoust. Soc. Am.* **120**, 3364–3364.

Fritz, C., Cross, I., Moore, B. C. J., and Woodhouse, J. (2007). “Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin,” *J. Acoust. Soc. Am.* **122**, 3640–3650.

Garcia, G. (2002). “Optimal filter partition for efficient convolution with short input/output delay,” in *Proc. of the 113th AES Convention*, Los Angeles, pp. 5660–5666.

Gardner, W. (1995). “Efficient convolution without input-output delay,” *J. Audio Eng. Soc.* **3**, 127–136.

Holm, J.-M., and Välimäki, V. (2000). “Modeling and modification of violin body modes for sound synthesis,” in *Proc. European Signal Processing Conf.*, Tampere, Finland, pp. 2229–2232.

Jansson, E. V. (2002). “Acoustics for violin and guitar makers,” <http://www.speech.kth.se/music/acvigit4/> (last viewed 6 November 6, 2010).

Jansson, E. V., Molin, N.-E., and Saldner, H. O. (1994). “On eigenmodes of the violin—electronic holography and admittance measurements,” *J. Acoust. Soc. Am.* **95**, 1100–1105.

- Jansson, E. V., Molin, N.-E., and Sundin, H. (1970). "Resonances of a violin body studied by hologram interferometry and acoustical methods," *Phys. Scr.* **2**, 243–256.
- Karjalainen, M., Välimäki, V., Penttinen, H., and Saastamoinen, H. (2000). "DSP equalization of electret film pickup for the acoustic guitar," *J. Audio Eng. Soc.* **48**, 1183–1193.
- Maestre, E., Bonada, J., Blaauw, M., Pérez, A., and Guaus, E. (2007). "Acquisition of violin instrumental gestures using a commercial EMF device," in *Proc. Int. Computer Music Conf.*, Copenhagen, Denmark, pp. 386–393.
- Marshall, K. D. (1985). "Modal analysis of a violin," *J. Acoust. Soc. Am.* **77**, 695–709.
- Morset, L. H. (2001). "A low-cost PC-based tool for violin acoustics measurements," in *Proc. Int. Symp. Musical Acoust.*, Perugia, Italy, pp. 627–631.
- Oppenheim, A. V., and Schaffer, R. W. (1975). *Digital Signal Processing*, (Prentice Hall Press, Upper Saddle River, NJ), Chap. 10, pp. 480–531.
- Otčenášek, Z., and Syrový, V. (1999). "Directivity of the violin radiation," *J. Acoust. Soc. Am.* **105**, 1330–1330.
- Pätynen, J., Pulkki, V., and Lokki, T. (2008). "Anechoic recording system for symphony orchestra," *Acust. Acta Acust.* **94**, 856–865.
- Perez, A. (2009). "Enhancing spectral synthesis techniques with performance gestures using the violin as a case study," Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain.
- Schleske, M. (2002). "Empirical tools in contemporary violin making," *J. Catgut Acoust. Soc.* **4**, 43–61.
- Smith, J. O. (1983). "Techniques for digital filter design and system identification with application to the violin," Ph.D. thesis, CCRMA, Stanford University, Stanford, CA.
- Smith, S. W. (1997). *The Scientist and Engineer's Guide to Digital Signal Processing* (California Technical Publishing, San Diego, CA), Chap. 17, pp. 297–310.
- Vos, H., Warusfel, O., Misdariis, N., and de Vries, D. (2003). "Analysis and reproduction of the frequency spectrum and directivity of a violin," *J. Acoust. Soc. Netherlands* **167**, 1–11.
- Wang, L. M. (2000). "Radiation mechanisms from bowed violins," Ph.D. thesis, Pennsylvania State University, Williamsport, PA.
- Wang, L. M., and Burroughs, C. B. (2001). "Acoustic radiation from bowed violins," *J. Acoust. Soc. Am.* **110**, 543–555.
- Waxman, J. (2005). "Violin acoustic radiation synthesis: A source model for direct sound enhancement in musical acoustic environments," *J. Undergrad. Res.*, University of Rochester **4**(1), 25–28.
- Weinreich, G. (1997). "Directional tone color," *J. Acoust. Soc. Am.* **101**, 2338–2346.
- Weinreich, G., and Arnold, E. B. (1980). "Method for measuring acoustic radiation fields," *J. Acoust. Soc. Am.* **68**, 404–411.
- Williams, E. G., Maynard, J. D., and Skudrzyk, E. (1980). "Sound source reconstructions using a microphone array," *J. Acoust. Soc. Am.* **68**, 340–344.
- Woodhouse, J. (2005). "On the 'bridge hill' of the violin," *Acust. Acta Acust.* **91**, 155–165.