Supplementary Material

Precomputed Wave Simulation for Real-Time Sound Propagation of Dynamic Sources in Complex Scenes

Nikunj Raghuvanshi, John Snyder, Ravish Mehra, Ming Lin, Naga Govindaraju

Error Analysis in Living Room Figures 1, 2, and 3 analyze error at three different listener locations in the furnished living room scene. See also Figure 7 in the main paper, which plots error at a fourth listener location. Error is shown in two parts: first (top of each figure), from encoding alone using our representation of peak times/amplitudes and a frequency trend, and second (bottom of each figure), from both encoding and interpolation at a listener location midway between the simulated grid points. Errors are computed with respect to a higher-frequency (4kHz) wave-based reference simulation, and so include frequencies beyond our "working" simulation which is bandlimited to 1kHz. Overall, compression error is low, often near the threshold of audibility, while total error (interpolation + compression) is somewhat higher but still reasonably small. In all cases, our method of interpolating peak times and amplitudes better preserves the high-frequency content of the impulse response than does straightforward linear interpolation of the signals. Not only is our result more accurate, but it also avoids "gurgling" artifacts from linear interpolation in which high sound frequencies are alternately preserved at the grid points and then attenuated between them, as the listener or sources move.

"Shoebox" Experimental Results We have implemented the Image Source method as a second reference solution, based on a simple, rectangularly-shaped room. The walls are assumed to be purely specular and without frequency-dependent absorption. Frequency-dependent atmospheric absorption is however taken into account. The main paper mentions that above 5kHz, frequencies are strongly attenuated when propagating in air. The spectrogram on the right of Figure 4 demonstrates this well-known result.

Figure 5 compares errors between the frequency response obtained with the Image Source method and our approximation based on a numerical simulation limited to 1kHz, encoded by peak delays and amplitudes plus a frequency trend, and interpolated spatially. In this simple scene, the Image Source method provides perfect interpolation, yielding the response at any desired point and so serving as a good reference.[1] Our results agree well with this reference. While the maximum error is around 5 dB, the average error is roughly 2-3 dB over the whole spectrum up to 16kHz. These errors very closely match those obtained when comparing with the wave-based reference simulation. Linear interpolation performs far worse, underestimating energy in all octaves above 2kHz by about 7-10 dB.

---

[1] The corners of a room do produce diffracted scattering which is captured by our simulator but not the Image Source method. This can be safely disregarded by choosing high wall reflectivity and keeping the source and listener far from the corners, as we have done in this experiment.

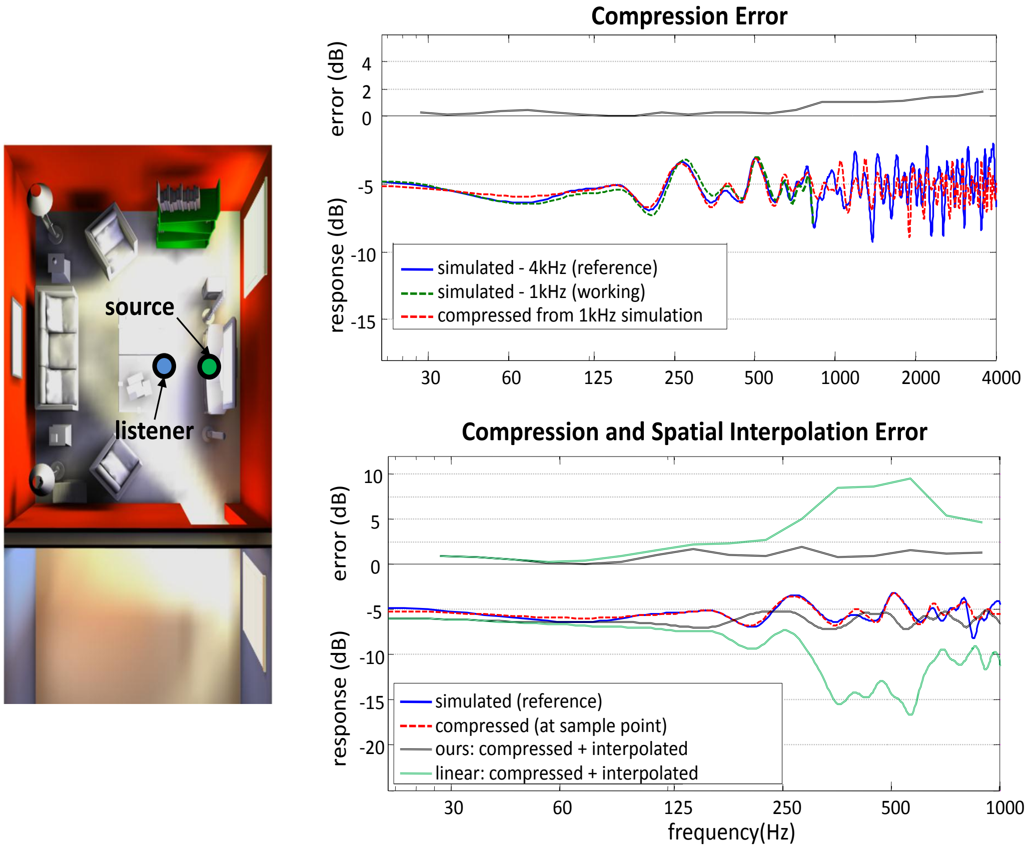

Figure 1: Error analysis compared to broadband numerical simulation: listener close to source. Our method matches the reference solution very closely, while linear interpolation yields substantial errors.

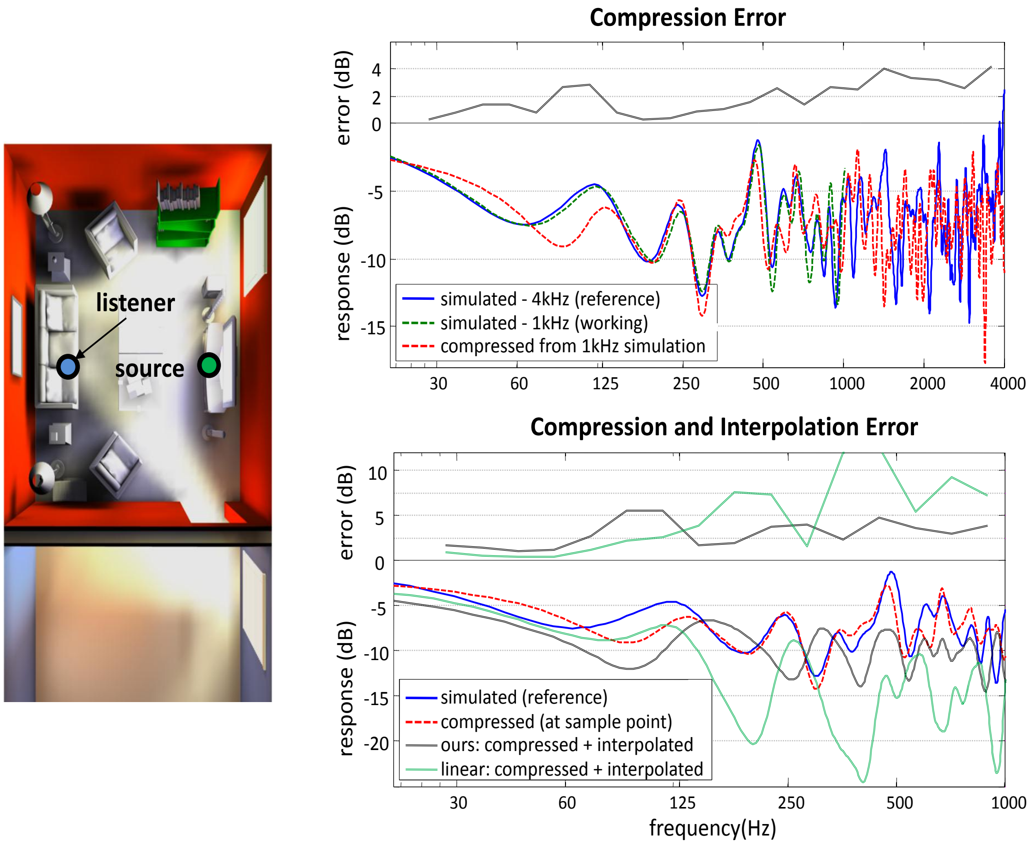

Figure 2: Error analysis compared to broadband numerical simulation: listener on couch. Compression error stays around 2dB till 1kHz and then increases to 4dB at 4kHz. Linear interpolation produces much more error.

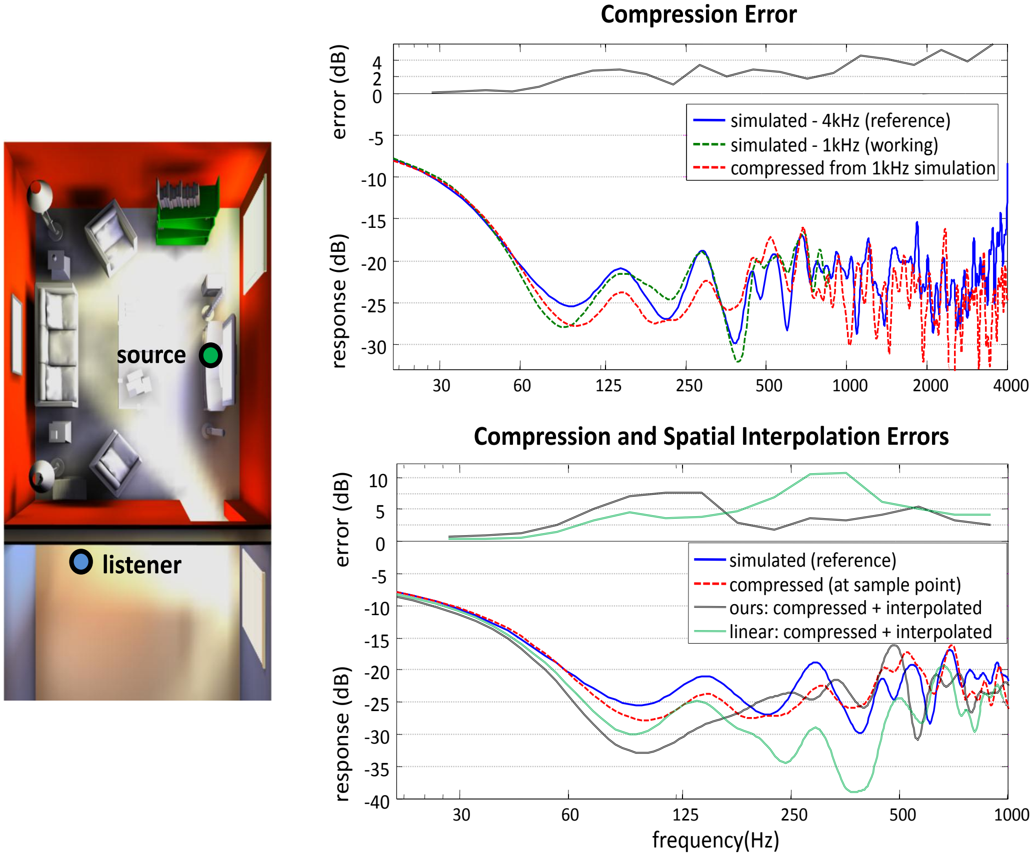

Figure 3: Error analysis compared to broadband numerical simulation: listener outside door. Sound from the source undergoes multiple scattering in the room, then diffracts around the door to arrive at the listener. Such high-order effects are very hard to model convincingly and a challenging case for current systems. Notice the clear low-pass filtering in the frequency response plotted and audible in the demo. Compression error lies between 2 to 4 dB, which is quite low. Spatial interpolation errors are higher, crossing 5 dB, but our technique produces less error over all frequencies than linear interpolation.

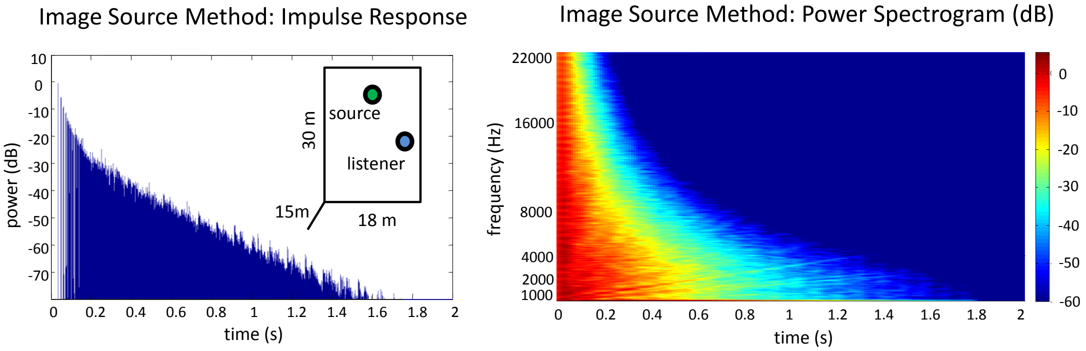

Figure 4: Effect of atmospheric attenuation. Results for the Image Source method on a simple "shoebox" geometry shown in the upper-right inset are plotted on the left. Impulse responses were calculated assuming a frequency-independent pressure absorption coefficient of 0.15 at the room walls, but with frequency-dependent atmospheric attenuation using the formulae given in ISO 9613-1 assuming an air temperature of 20°C and relative humidity of 40%. The spectrogram on the right shows strong attenuation of frequencies above 5kHz. In particular, a 10kHz sound component decays by nearly 20 dB after just 200 ms of propagation. In a real scene where frequency-dependent material absorption is also accounted for, the attenuation would be even higher.

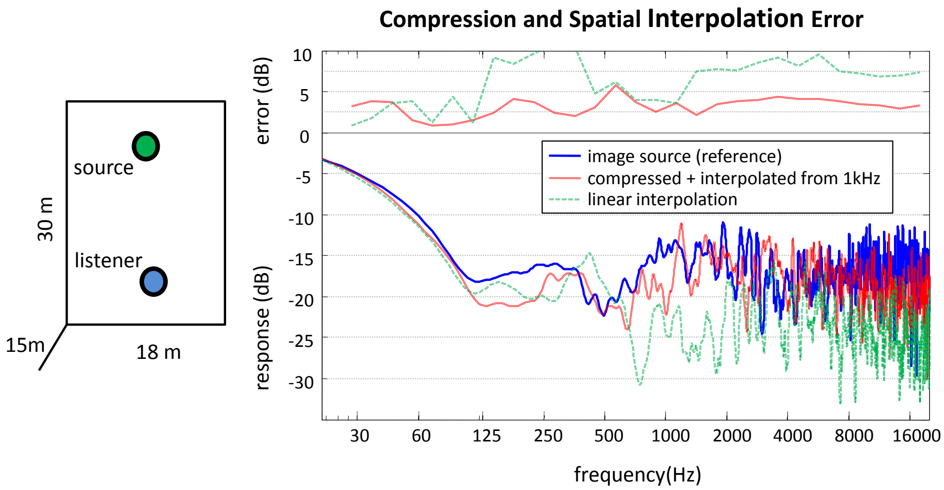

Figure 5: Error analysis compared with Image Source method. The left image shows the locations of source and listener in a simple "shoebox" room. A reference 200 ms long IR was generated with the Image Source method. The resulting frequency response is compared with our result based on a wave simulation bandlimited to 1kHz, and including compression and spatial interpolation errors. Our result (red curve) agrees well with the reference solution. In the top error plot, maximum error is about 5 dB while average error over the full frequency range up to 16kHz is 2-3 dB. Linear interpolation (green curve) yields larger errors and incorrectly attenuates higher frequencies.