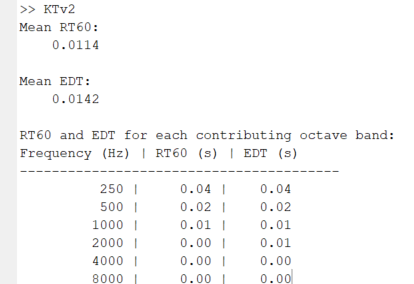
# Monday, 19 August 2024

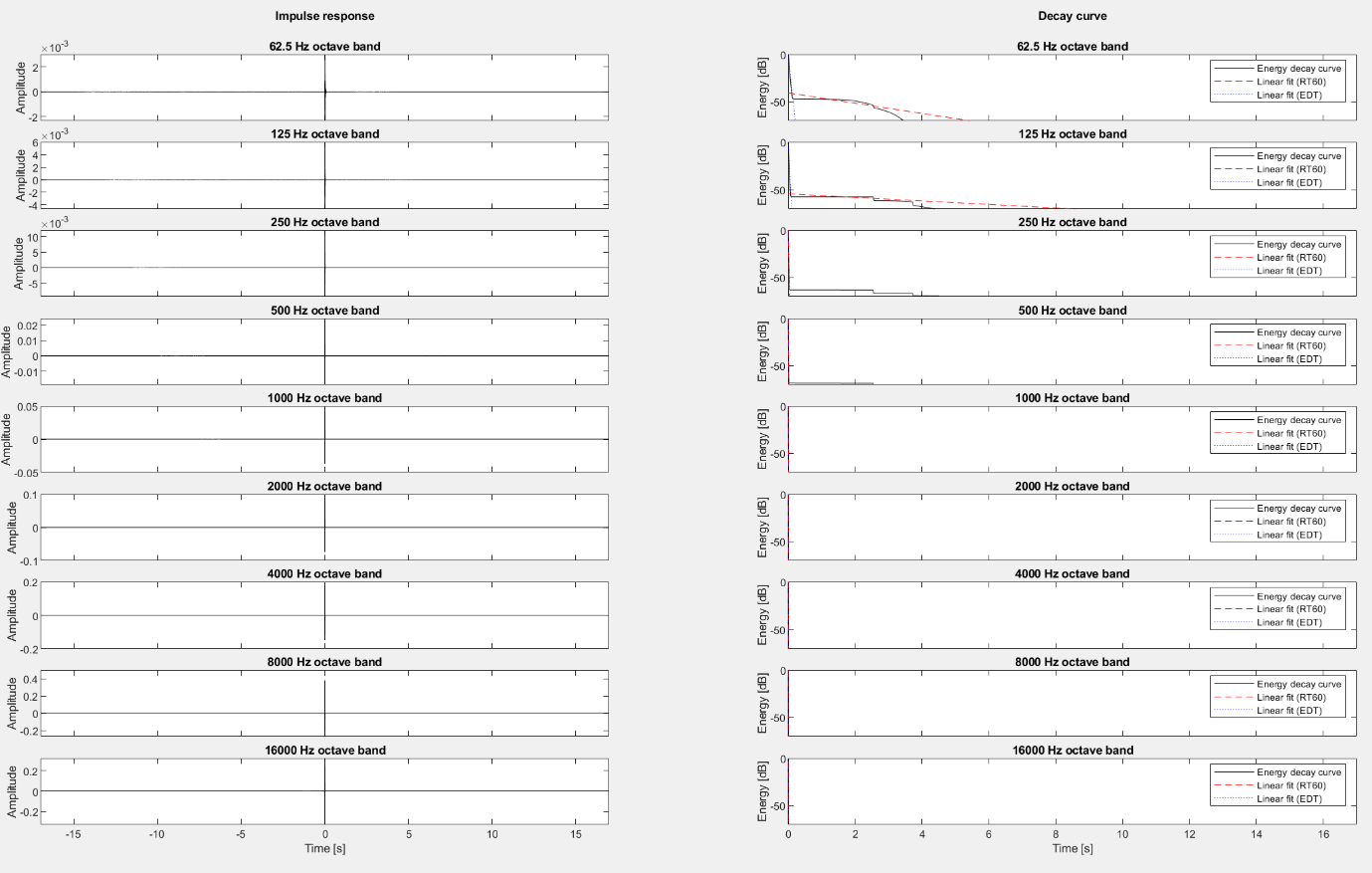
WFH, restudy RIR and deconvolution. Also looked into existing Unity project that have RIR/realistic simulation (not found).

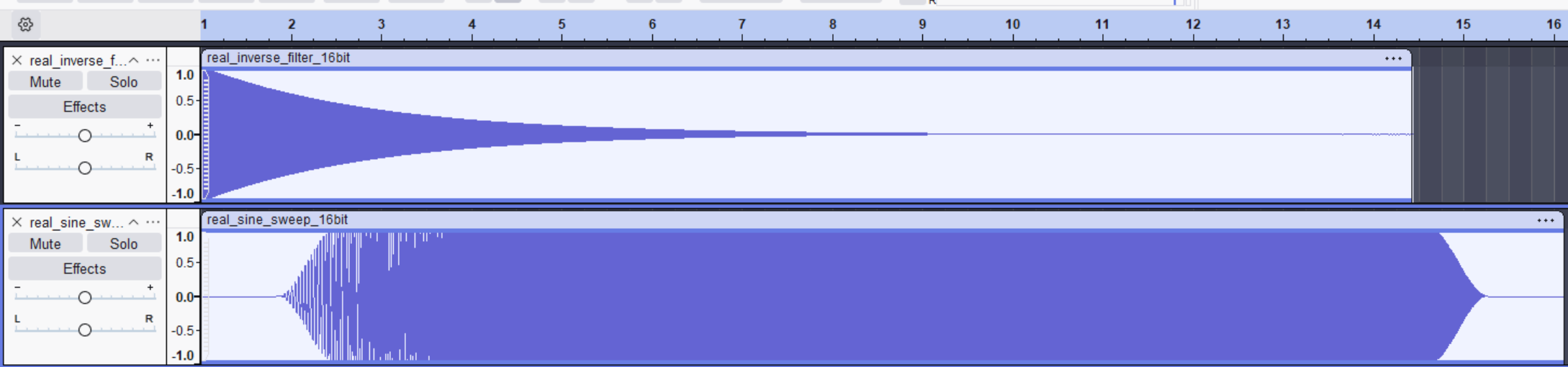
Tested RIR using generated sweep as recorded sweep:



Im assuming this anomaly (floor) is due to inaccurate inverse filter thus causing problem after deconvolve.

Instead of generating the sine sweep and inverse filter ourselves, let’s try using audacity plugin instead.



Old sine sweep and inverse filter: 

# Tuesday, 20 August 2024

## Meeting w/ Dr Hansung Kim:

### What I did last 2 weeks:

Implemented direct Unity audio recording  
Created test scenes to isolate Steam Audio issues  
Experimented with various audio parameters

### Problems encountered:

Deconvolution/inverse filter inaccuracy issue.  
Audio clipping limiting further tuning  
High EDT value (compared to RT60)  
Peak in open air test environment

### Plans:

Regenerate sine sweep and inverse filter using audacity plugin instead of python manual coding to eliminate errors etc  
Retest on test scene and generated as recorded to sanity check  
Have a meeting with Mona for evaluation progress.  
Update progress with Dr, Mona and Atiyeh.

### Take time off on 29 and 30, but work on 2 September instead. Meeting w/ Dr on 2nd September, and meet Mona for system handoff (laptop, VR HMDs) on different date.

Make a demo video or try to make it easy for VR to work (whichever most convenient to demo)

Send poster for print before 28th.

Prepare report (around 20 pages), no deadline but don’t overwork outside intern time if possible.

Make it well structured for handoff! (Well documented and structured)

Think about writing papers as dual First author w/ Mona or secondary author (prob depends on how well I got result evaluation going).

Let’s rewrite all this into proper todo list.

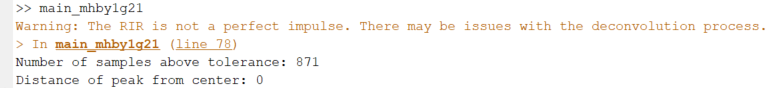
## Next 2 week TODO:

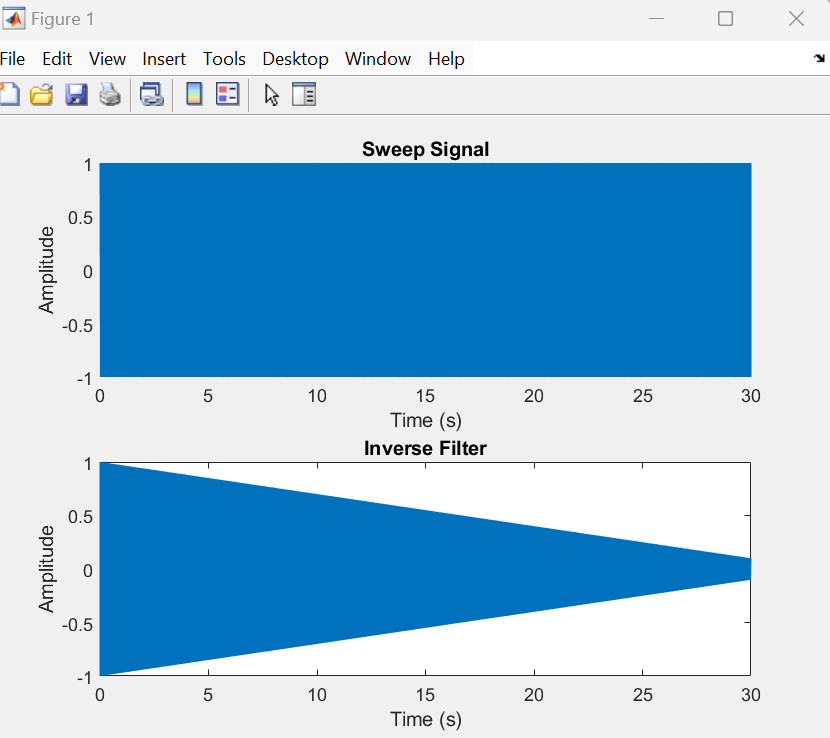
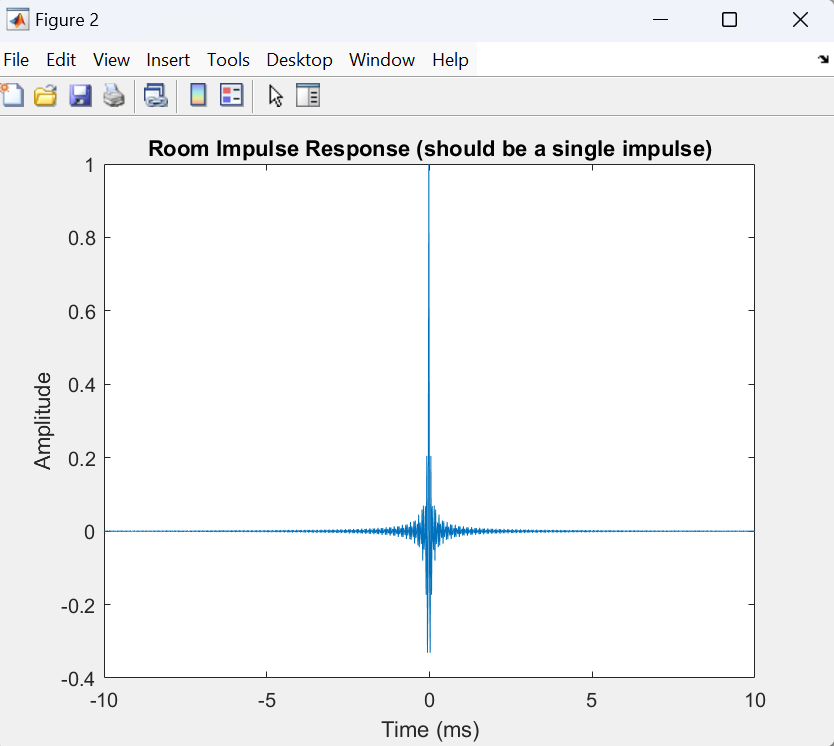
* Send poster for print before 28th.
* Make a demo video or try to make it easy for VR to work (whichever most convenient to demo)
* Prepare report (around 20 pages), no deadline but don’t overwork outside intern time if possible.
* Meeting w/ Dr Hansung on 2nd Septembe, Mona on any day after 3rd for handoff.
* Make it well structured for handoff! (Well documented and structured)
* Think about writing papers as dual First author w/ Mona or secondary author (prob depends on how well I got result evaluation going).

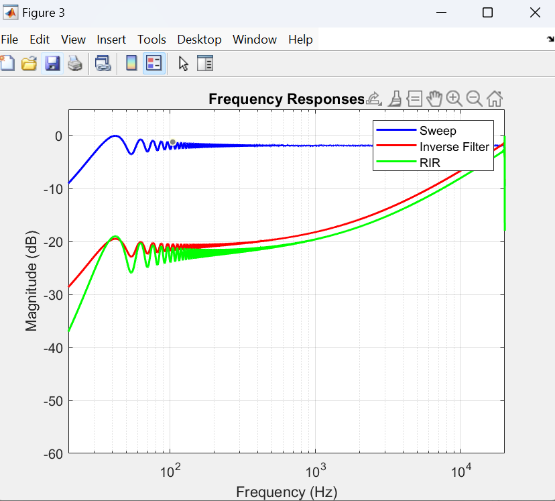
## Immediate TODO:

1. Regenerate sine sweep and inverse filter using audacity plugin instead of python manual coding to eliminate errors etc
2. Retest on test scene and generated as recorded to sanity check
3. Have a meeting with Mona for evaluation progress.
4. Update progress with Dr, Mona and Atiyeh.

Gave up on no. 1 because I think its impossible to get better approximate of sine sweep from input only, and generating sine sweep together with inverse filter (where parameter is known and defined) is a better way. Although because this would be exactly the same as Mona’s previous one, so instead I used matlab to generate the sine sweep and inverse filter instead of Python.

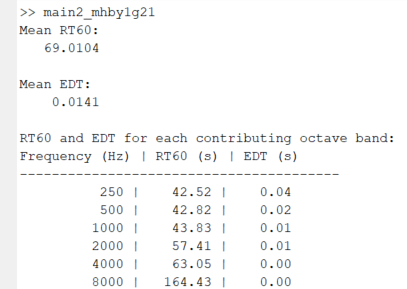
I also ran the deconvolve on the on the generated sweep straight away (without any simulation/noise) to get as perfect impulse response as I can, it still not perfect but good enough imo which prove the deconvolve is working. 

🡨 It looks a lot different than previously because its on linear instead of exponential I think.

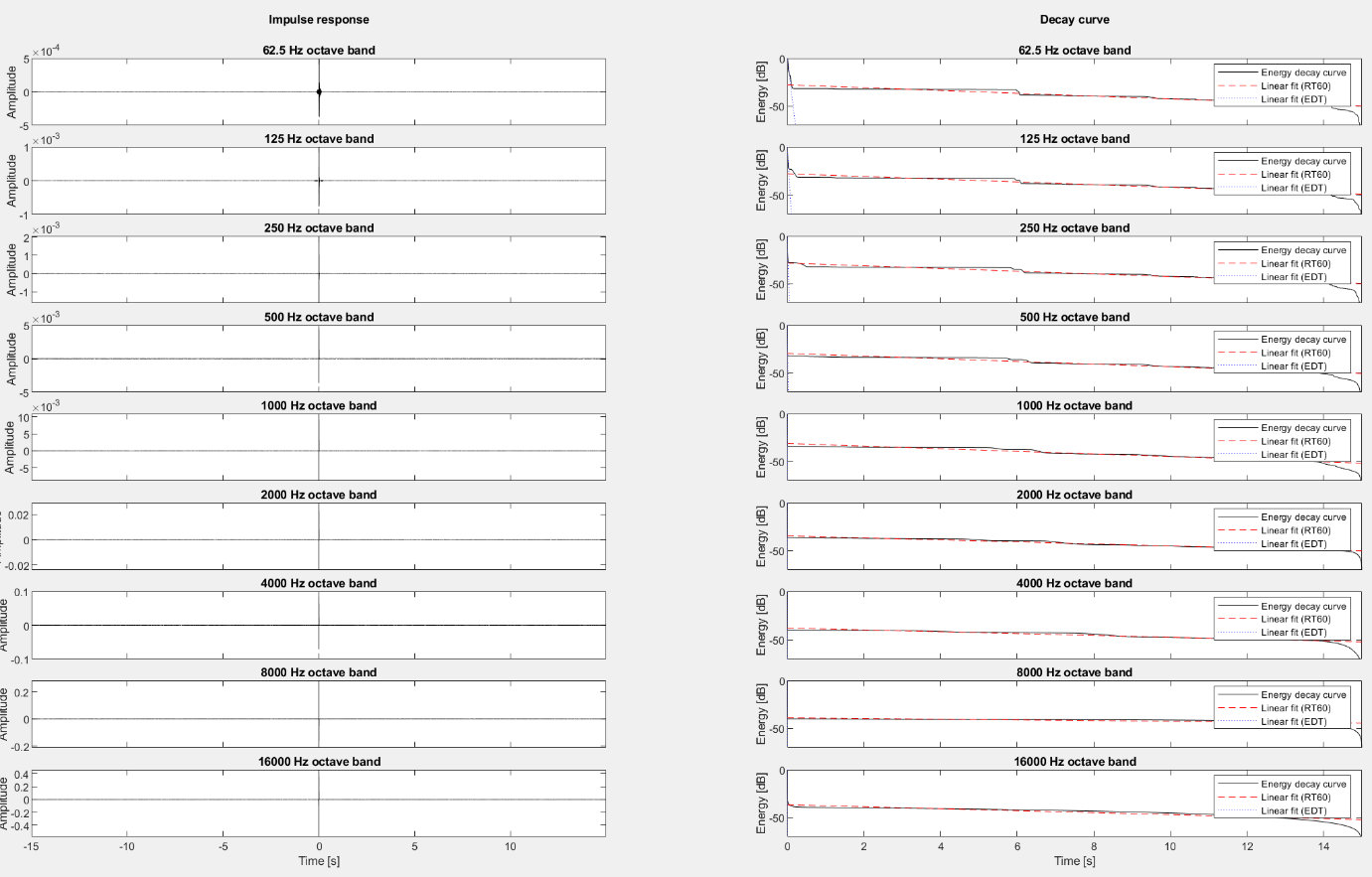
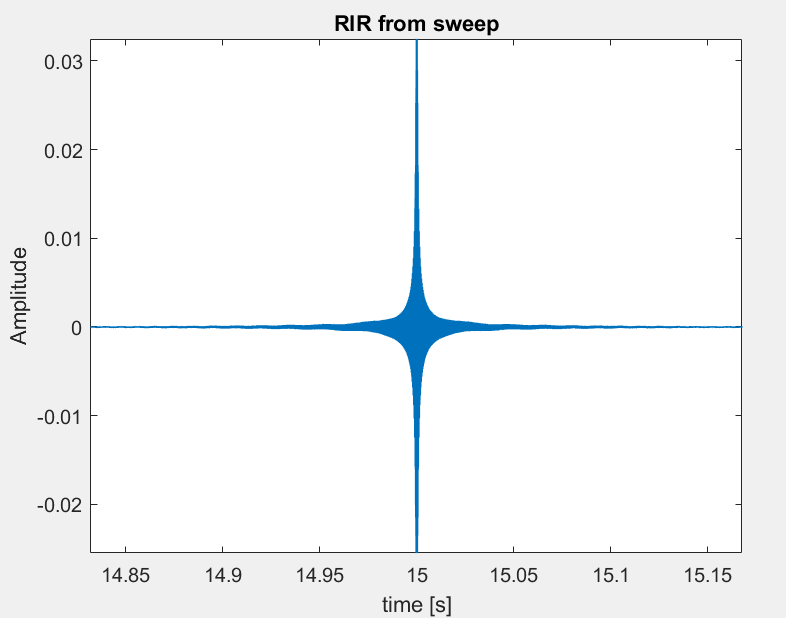
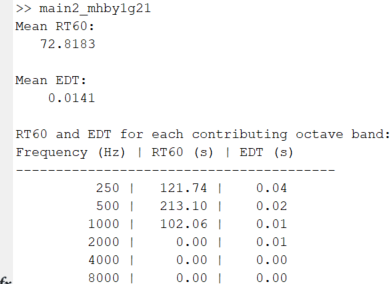
Notice the RIR is very short which is good sign (ms unit)

Also added Frequency response comparison for all 3 signals for sanity check according to this [stack exchange](https://dsp.stackexchange.com/questions/41696/calculating-the-inverse-filter-for-the-exponential-sine-sweep-method), which is different but im assuming this is because im using linear (need to revise later if this still gives problem).

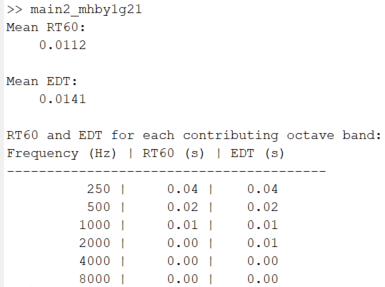
Let’s also add in the RT60 and EDT calculation to make sure.

Separated the 🡪

main\_mhby1g21.m to main1 and main2 respectively in order of to run them (will require three next for generation, rir deconvolve, and rt60/edt analysis). As seen on the right, its not working on default y\_fit due to noise floor again interestingly even though the RIR is really short and looks correct.

This means either my inverse filter or deconvolve is still wrong or maybe because im using linear, or noise floor is just inevitable, lets use y\_fit [-5 -35] first just to be sure.

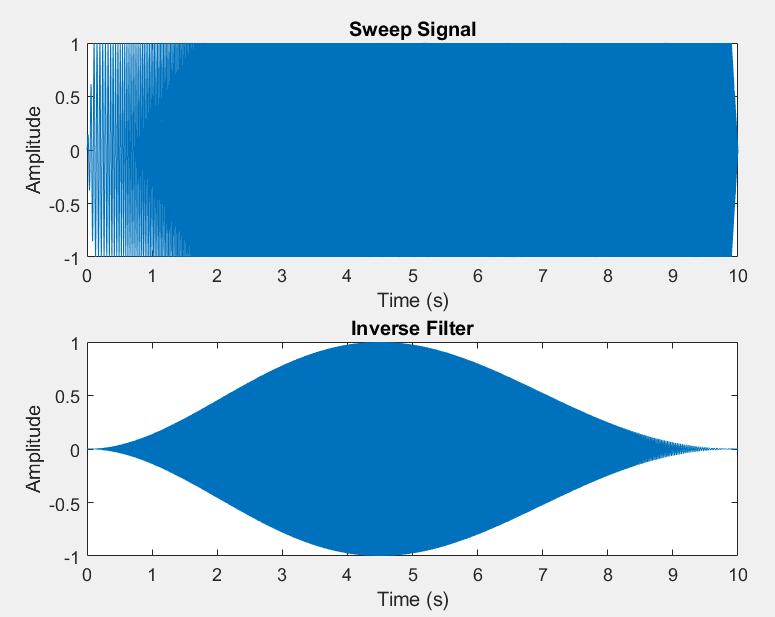
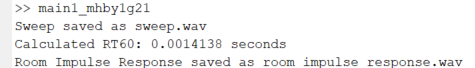
As seen on right, even -5 -35 is not enough due to noise floor anomaly.

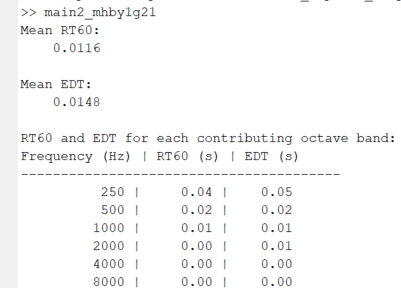
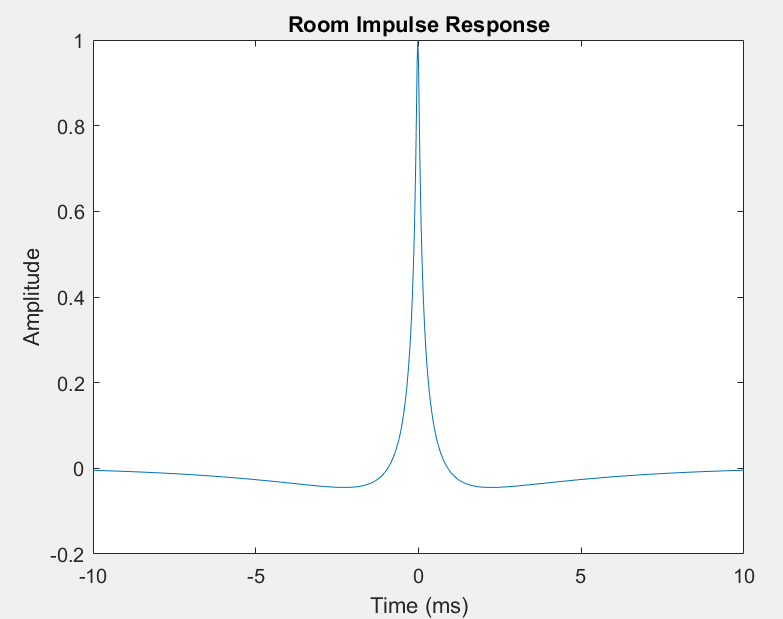
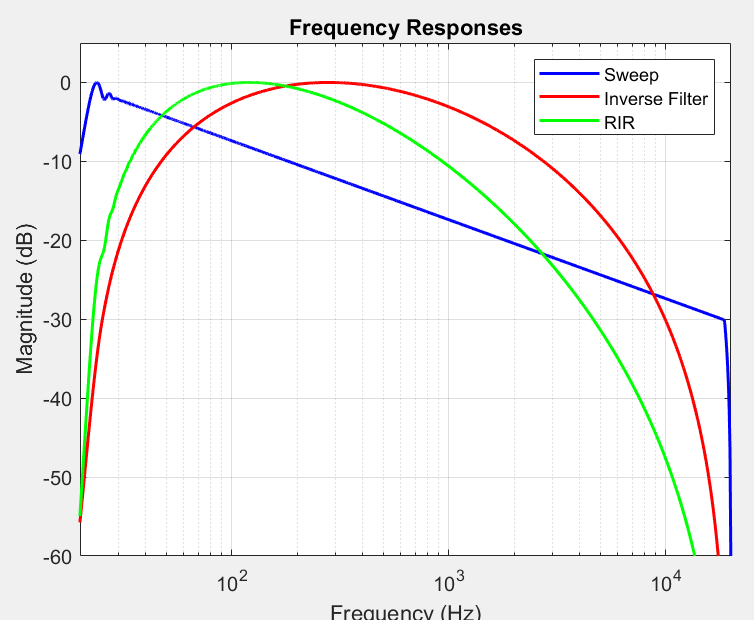
🡨-5 -25 give correct value though, however this is obviously not fit (pun intended) for use at all as this is the exact same sweep, not even simulated and it already have such high noise floor and anomaly/issues..

Ok interestingly enough, decreasing the duration of sine sweep lower the noise floor level but didn’t completely remove it, this means the problem might stem from floating point precision error cumulation or something like that.

Updated main1\_mhby1g21.m script with following improvements:

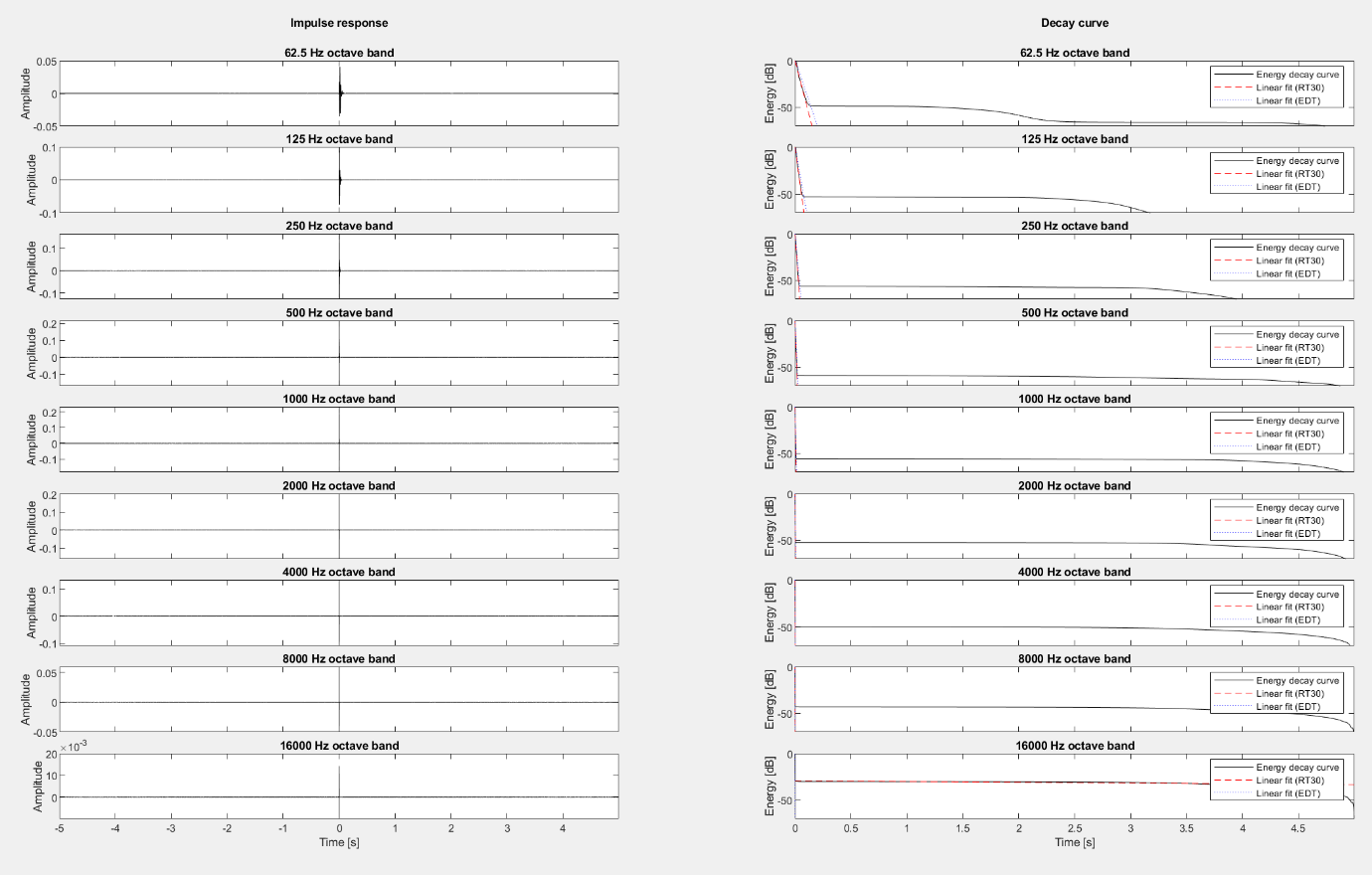
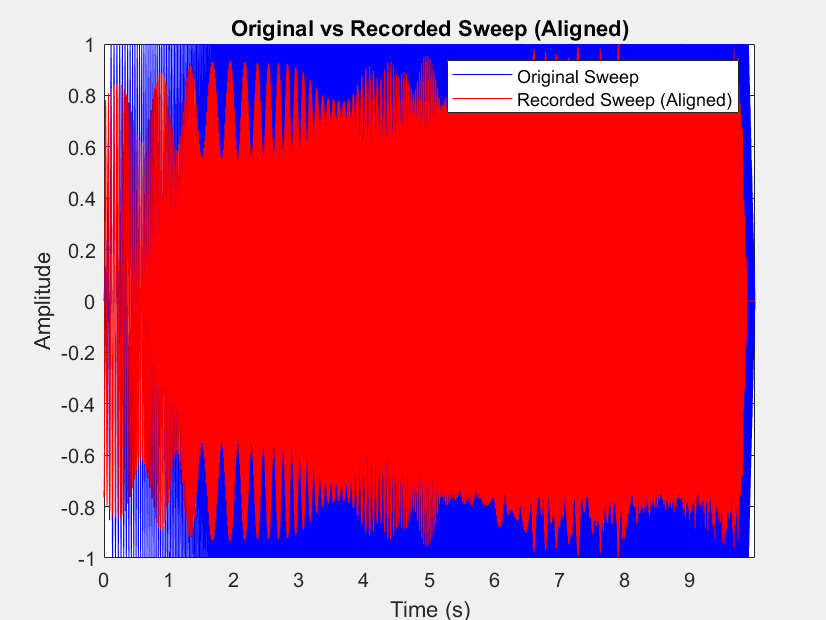
* Uses an exponential sine sweep instead of a linear chirp.
* Applies a fade-in and fade-out to the sweep.
* Uses a window function (Hann window) on the inverse filter.
* Ensures double precision for all calculations.
* Implements a custom RT60 calculation function with a noise floor cutoff.

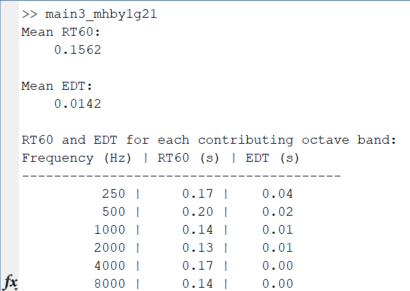
This results into best one so far! GOOD NEWS LETSGOO! 

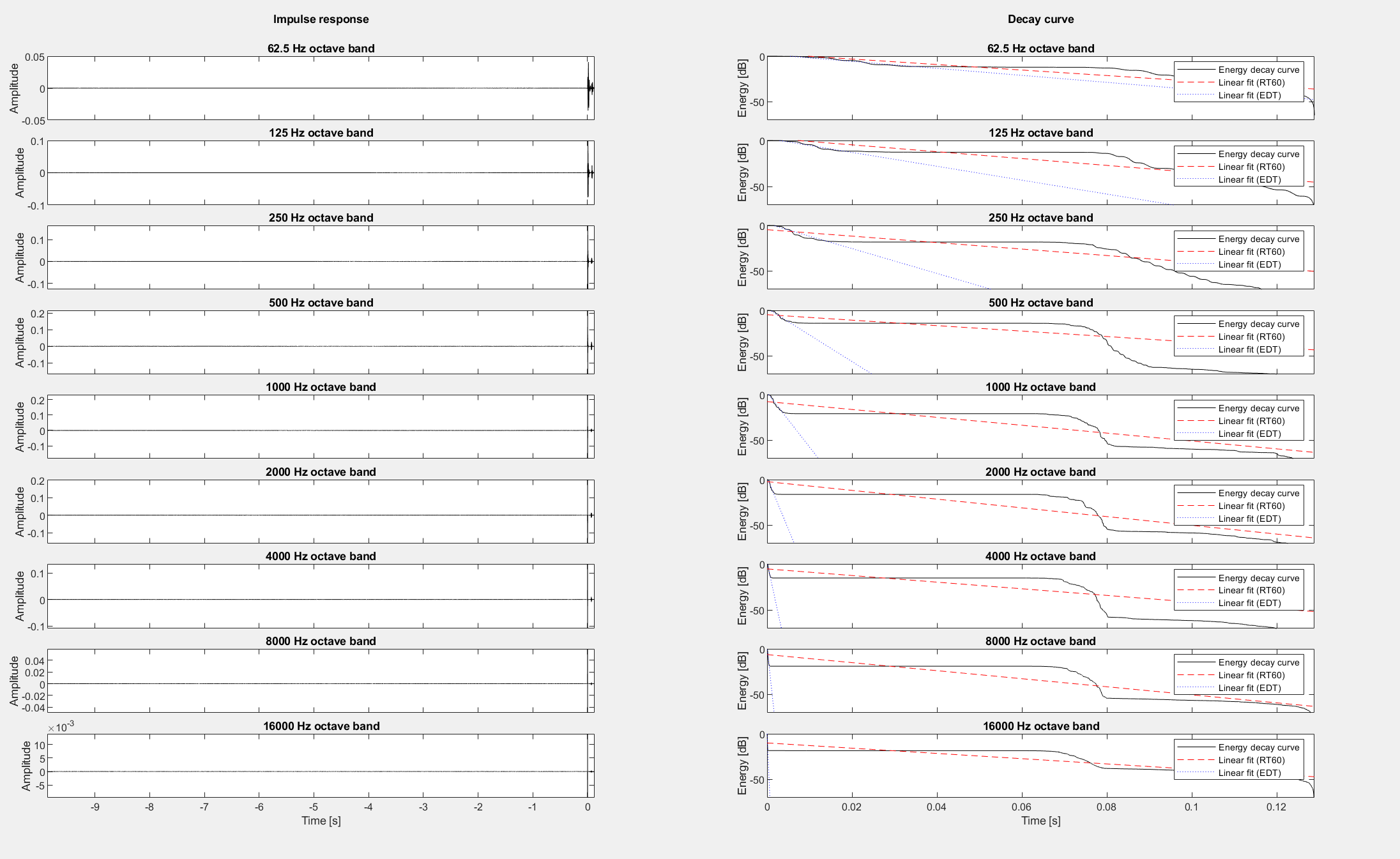


Look at how clean the RIR is (literally near dirac delta), idk what freq response really means here and the inverse filter/ sweep looks a bit strange due to modification but it seems working. This is obviously with recorded sweep being the generated sweep for sanity check.

As seen below, the noise floor still exist but atleast its way lower now that y\_fit [-5 -35] can be used reliably. Now lets run this on Unity and hope it works as well.

OK, it doesn’t work straight away obviously, theres lot variable in play especially in aligning the signals, worse part is that removing silence does not really work imo because certain frequency (esp lowest would probably not be captured in Unity, and thus the recorded sweep would seem shorter than it should) but this is the best I got for now (hack-ish cross correlation after silence removal. 

This looks correct but when u look at actual octave bands decay curve, it still have weird bump so idk.



We can change the silence threshold though so lets try that. Im gonna commit first just in case.

OK Nvm, changing the silence threshold from 0.001% to 0.0001% causes RT60 calc to gives lots of insufficient dynamic range warning and also give worse RT60 value due to higher floor somehow.

So the choice now imo is to modify our unity audio capture to automatically capture as soon as the audio is played so no need to realign on post.

Before doing that, lets try the working ish one (on top, 0.001% silence threshold) with KT scene just for sanity check. TBF, the result we got from open air is kinda similar ish to LR in Kim21 which should be the baseline imo.

Ok vol0.5 is too much for KT, lots of audio clipping and presumably also what causing not enough dynamic range warning.

Unfortunately going to vol0.1 also decrease RT60 as expected and didn’t resolve EDT issue (EDT higher than RT60)

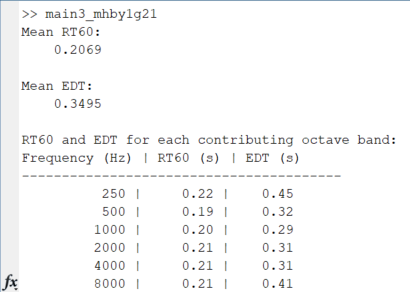
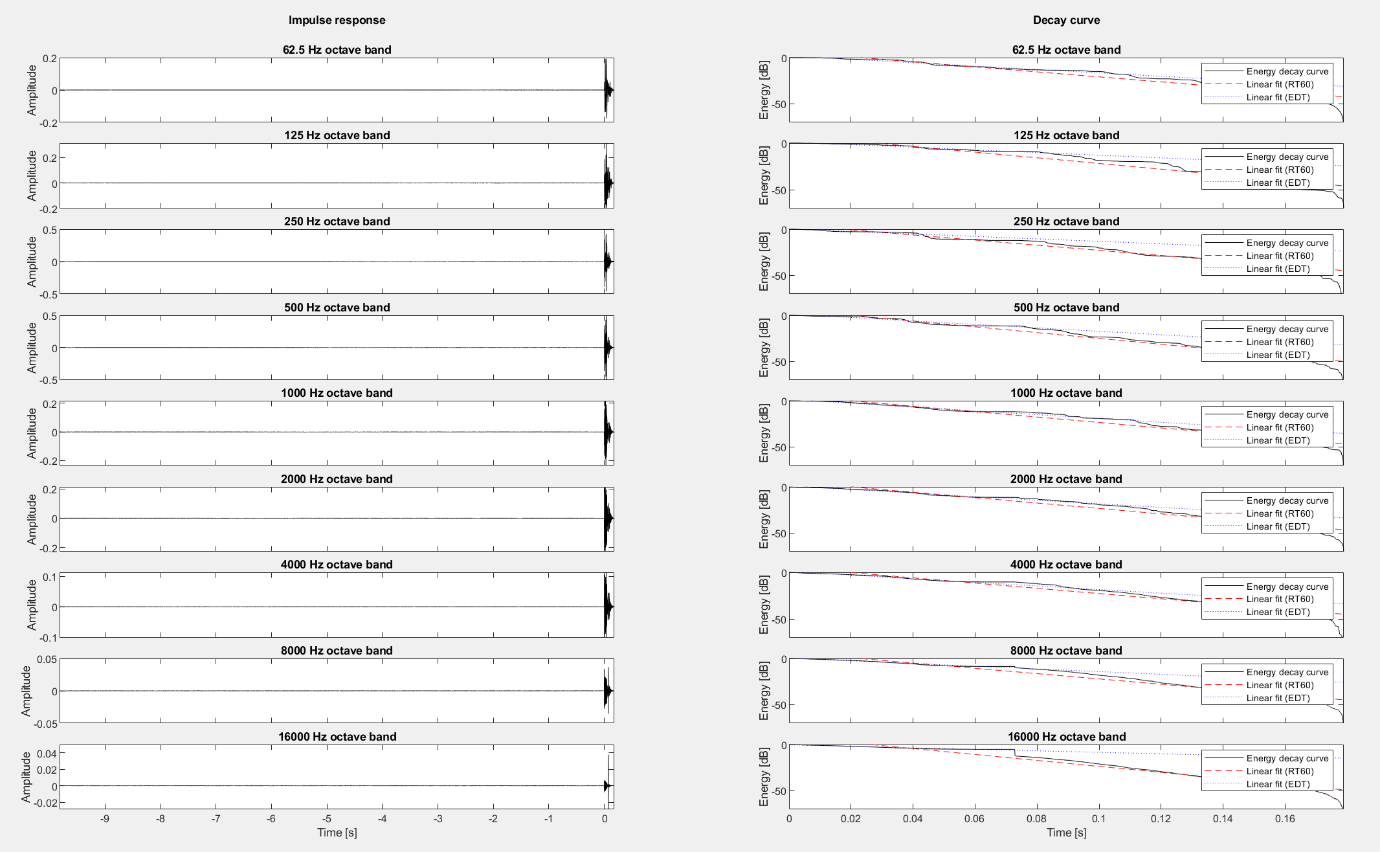
Using attenuation combined with 0.5to5 ratio doesn’t fix the issue either tho the former does allow us to use higher vol with less clipping to get more dynamic range.  
I’m convinced the issue for EDT might either be intrinsic with my KT scene or with the alignment, so alignment would be my last try, lets do that tomorrow first thing!

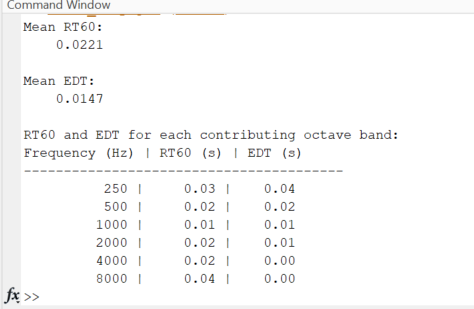
### Update to TODO:

* Sync Unity audio recording with audio player so no need to align in post.
* Retest on generated as recorded to sanity check, open air test, and all scene after.
* Have a meeting with Mona for evaluation progress.Update progress with Dr Hansung, Mona and Atiyeh.

# Wednesday, 21 August 2024

As said yesterday, lets sync audio recording with audio player. This is fairly simple, just modified AudioRecorder.cs a bit and added few lines on DelayedAudioPlay.cs to invoke the recording before audio start playing at exact same second.

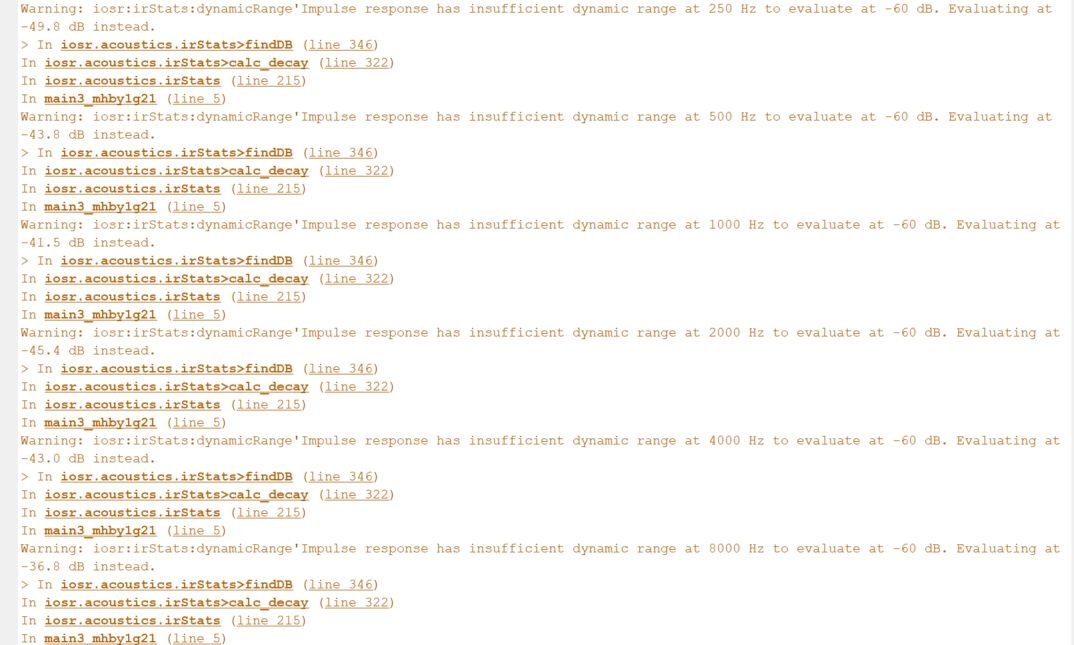
* 🡨 KT result, I’ve tried more modification on main2 for postprocess etc to try and get RT60 higher than EDT but this is to no avail. Atleast this one which have 2nd peak window remove and lots other preprocess don’t have noise floor level issue anymore..  
  Tested on LR and got very high RT and EDT, realized it was probably because silence removal messing

up the deconvolution, so changed from 0.001 to 0.0000001%. Here is the result on Test openair 🡨synced

OK, that percentage however is too low for LR, it causes low dynamic range warning so ig we need to fine tune each scene for different silence percentage..

NVM, that was not the problem, even on 0.01%, im still getting the error, must be due to the processing.. lets go back to simplest no processing like in main1 then…

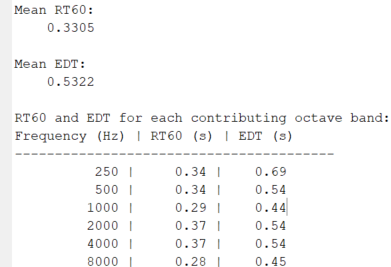
Interesting, using bare bones main2 with no preprocess like in main1 (perfect impulse response), for open air, im getting a second peak that is messing up the rt60 calculation but the edt is perfect. This only happen in open air test scene and KT scene remains the same as all before (higher edt than rt60 with same range of values)   
  
ok now after I look into LR scene again, it makes sense why mine is so high, the predicted material for mine is metal which is quite reverberant (hard, reflect a lot). Totally don’t use my LR as baseline for calibration because it is not right in the first place.

After looking back into the low dynamic range warning for LR previously, it was using a lower y fit than what we need to use on default (-5 -35) anyway, so might as well use this one (below)

So reverting back to the main2 with lots of processing (including silence removal) so atleast we don’t have noise floor level problem.

Ah just realized the silence removal is what remove the noise floor. Thus instead I should just preprocess it so it remove the flat noise floor level on every octave band

Also, our main problem right now with most of these scenes evaluation is that the RIR don’t behave in expected way (the decay curve), as EDT have higher value than RT60 which suggest the shape of the curve is strange (it should’ve been exponential from start, ie steeper earlier instead of later)

Even when using 0.5to5 ratio and using silence removed main2, it still give such values… I am honestly really sstumped and stressed..  
I think best course of action now is to just modify main2 to postprocess the signal so that on each octave band, after the lowest -db had been reached, remove all signal after to clean up the noise floor.. or check if the signal go back up or stays at very low slope/gradient..

For now, lets email Mona for a meeting to discuss our finding and issues.

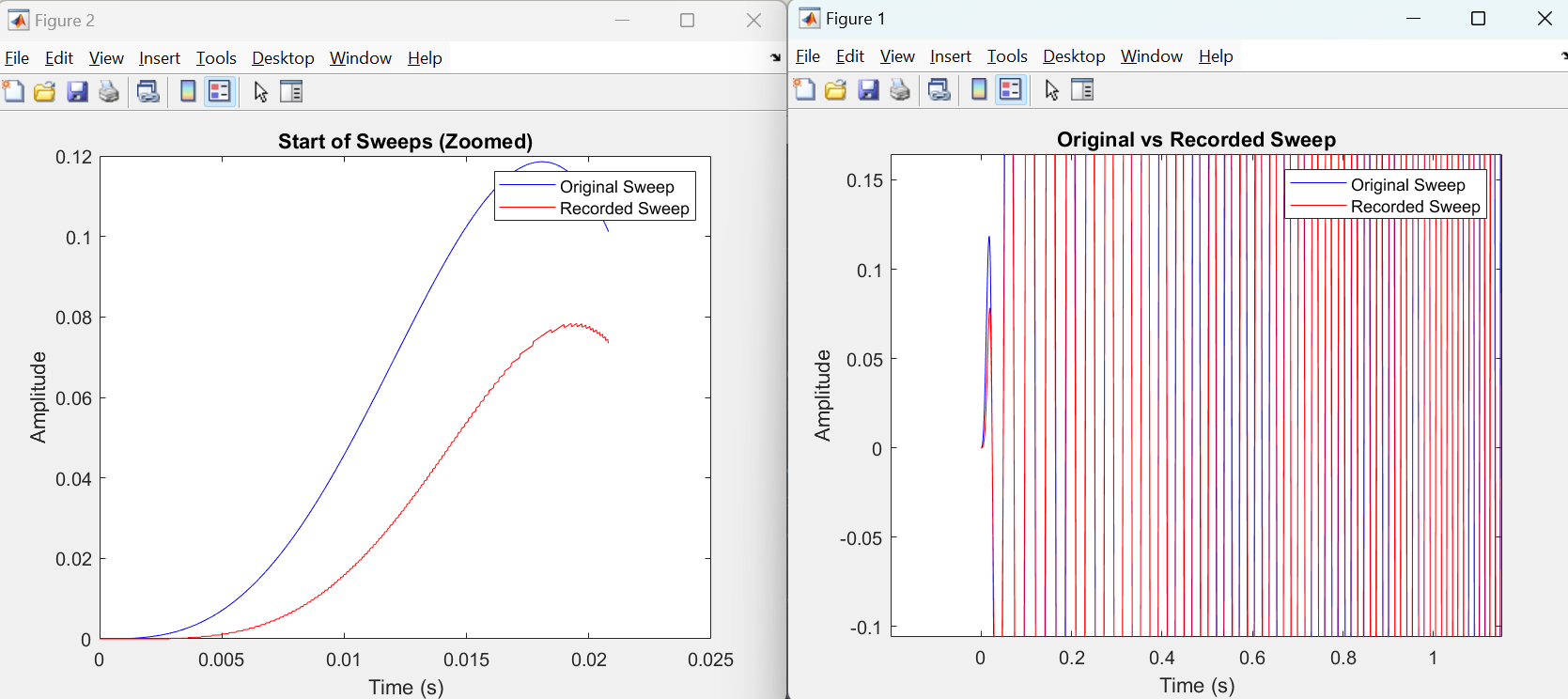
For tomorrow, lets use the silence threshold in main2 and generate all scenes to see what we got so far (with different silence threshold for each to remove noise floor). Personally I think the inverse filter/deconvolve method including align (thanks to sync implementation in unity) is perfect because our RIR looks pretty good, but would be great to reconfirm w/ Mona & Atiyeh.

# Thursday, 22 August 2024

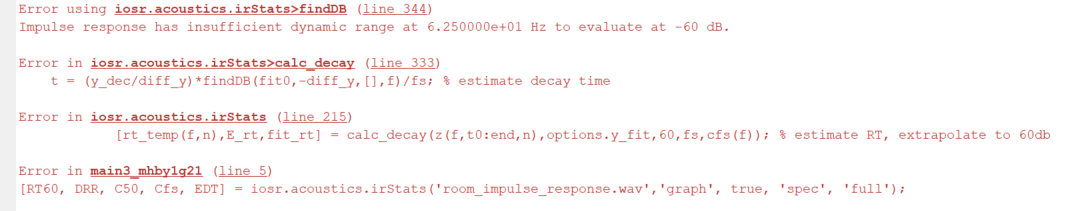
Mona replied with meeting set on Tuesday next week (27th), I however need to send my poster for printing before 28th, so lets just continue with generating all scenes result as planned and see how it is. Then just start creating poster and start wrapping up/writing technical report and readme for handoff etc.

OK last one modification before trying on all, lets realign the signals because I just realized its not as aligned as possible which might be culprit for the issues.

This is now the most aligned I can get (literally getting orange color from blue and red)

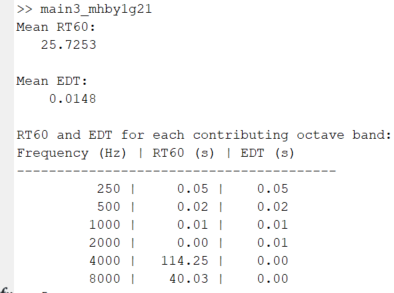


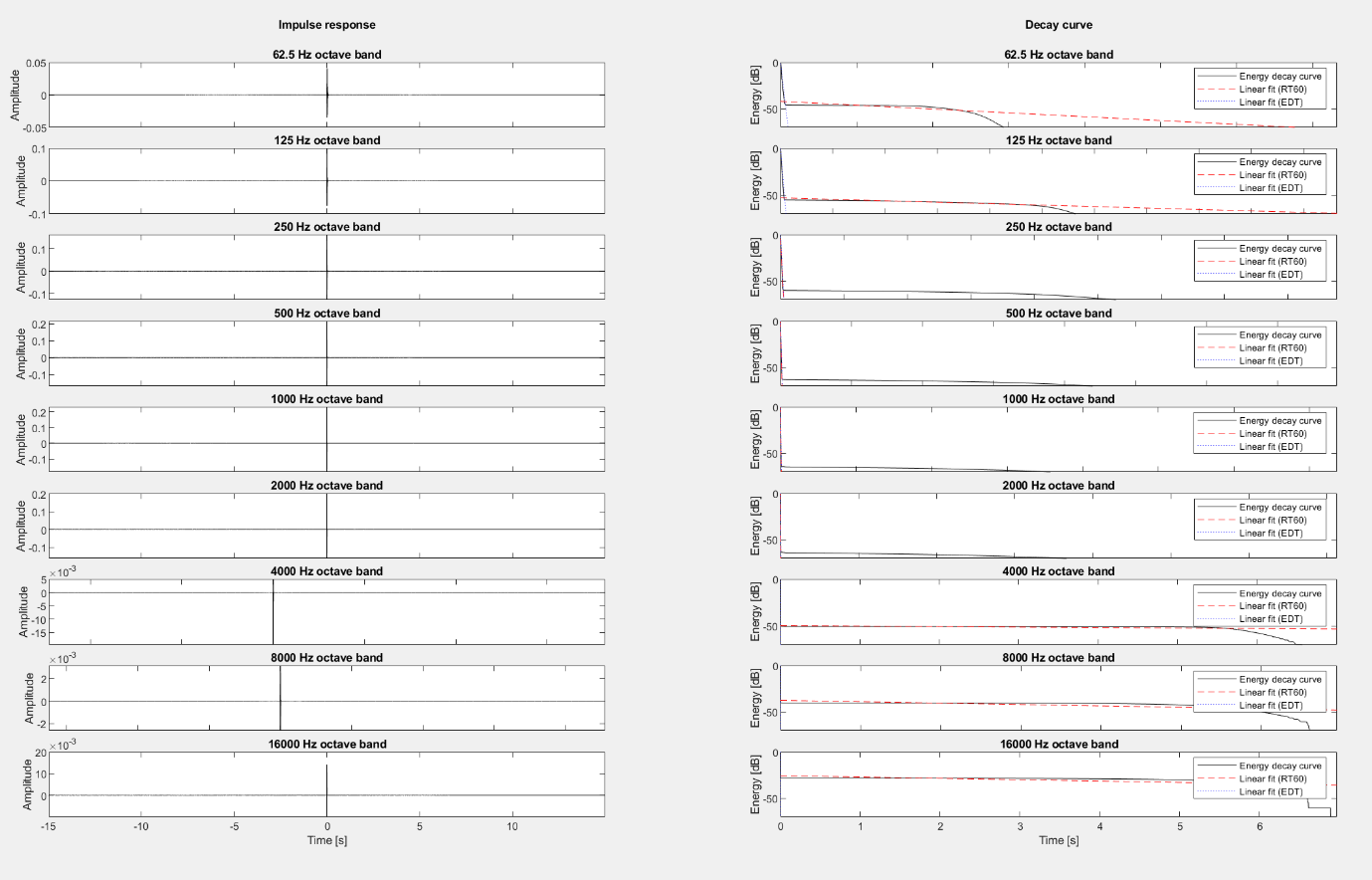
Unfortunately, for some reason, this causes error on main3.

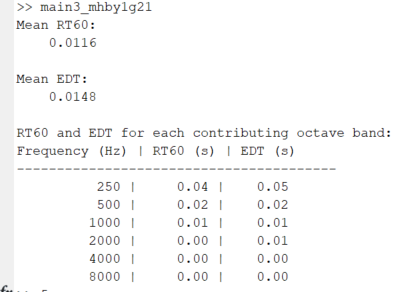


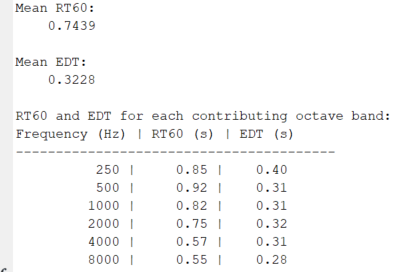
Nvm, this only perfectly align for open air test, KT isn’t align for some reason.

Added alignment window method, and this works pretty well for KT and LR imo but failed horribly on Test open air for some reason… For now lets get more recording (from ST, UL and MR for sanity check)

NVM, after rerunning it more, I feel like on lower freq, the rt60 and edt is higher than usual. Tbf this could just be because of the length of our sine sweep. So let’s try longer sine sweep next.  
Here is perfect impulse response on 30 sec on default y\_fit

As seen, the noise floor is still higher than it should be for higher and lower freq so lets use y\_fit [-5 -35] for these.

Much better!  
  
Lets get the response from open air, KT, LR, ST, UL, MR now.

Below is KT for 30sec sweep, so nvm, its not at all better.

I gave up for today.

# Friday, 23 August 2024

Before going back there, lets do some basic poster and report stuff.

Wrote basic headings and sections for report and created A1 poster pptx.

OK, lets go back to basic deconvolve with pure silence y=0 removed at start except last point

Also experimented with frequency domain manual multiplication (manual convolution) but no avail, same result as expected. And also realized the time sync is prob not the culprit as deconvolution (convolution) happens in frequency domain anyway…

Testing different setting and scene config (standardize distance and probe) for test open air scene because I still have no clue where the second peak coming from…still not fixed, probably a quirk with unity audio, hearing it subjectively, I don’t hear anything amiss either…

I gave up, lets just use basic deconvolve with default y\_fit and get result again for each scene.