



به نام خدا



دانشگاه تهران
دانشکده مهندسی برق و کامپیوتر
پردازش سیگنال‌های زمان-گسسته
تمرین کامپیوتری سری ۳

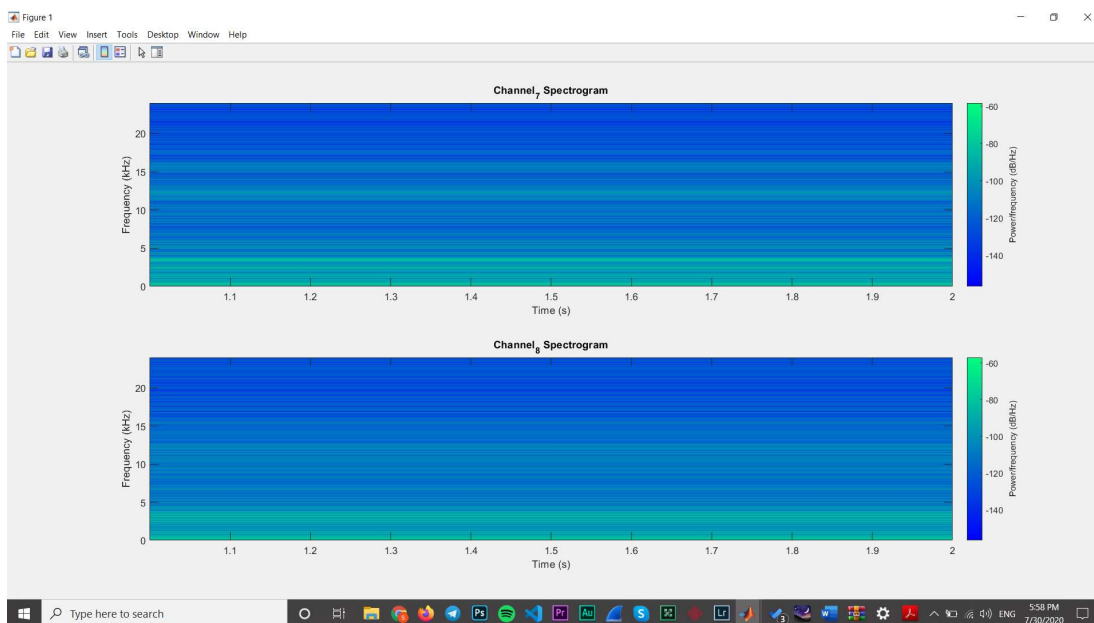
نام و نام خانوادگی	سهند خوشدل
شماره دانشجویی	۸۱۰۱۹۶۶۰۷
تاریخ ارسال گزارش	۹۹ / ۵ / ۹

فهرست سوالات

سوال ۱ -	۲
سوال ۲ -	۳
سوال ۳ -	۶
سوال ۴ و ۵ -	۷
سوال ۶ -	۹
سوال ۷-۱۰ -	۱۲

سوال ۱ -

Spectrogram نوعی نمایش اطلاعات زمان - فرکانس سیگنال است، که به کمک آن می توان محتوای فرکانسی سیگنال را در هر بازه زمانی مشاهده کرد. اسپکتروگرام معمولاً با نام های دیگری مثل **Voiceprints** یا **voicegrams** نیز مورد استفاده قرار می گیرد. از اسپکتروگرام می توان برای شناسایی کلمات گفته شده به صورت آوایی و تحلیل آواهای مختلف حیوانات استفاده کرد. شکل زیر نمایش اسپکتروگرام برای کانال های ۷ و ۸ را نشان می دهد. این شکل ها شباهت زیادی با هم دارند چرا که تنها اختلاف فاز ناشی از محل قرارگیری میکروفون در آن ها متفاوت است



سوال ۲ –

در این قسمت سیگنال های ورودی کانال های ۷ و ۸ را در حوزه فرکانس به چهار بخش تقسیم می کنیم

(این تقسیم بندی شامل یک بخش پایین گذر، دو بخش میان گذر و یک بخش بالا گذر می باشد.)

- برای این کار ۴ فیلتر FIR با متد equi-ripple طراحی شد که از مرتبه بالایی برخوردارند و با انتخاب بازه هایی به طول ۵۰ هرتز خول فرکانس های قطع برای F_{pass} , F_{stop} سعی شد بهترین عملکرد حاصل شود.

- Cut off Frequencies : 6000,12000,18000,24000 (Hz)

Low pass filter:

F_{pass} : 5950

F_{stop} : 6050

Band pass filter 1 :

F_{pass} : 6050 , 11950

F_{stop} : 5950, 12050

Band pass filter 2 :

F_{pass} : 12050 , 17950

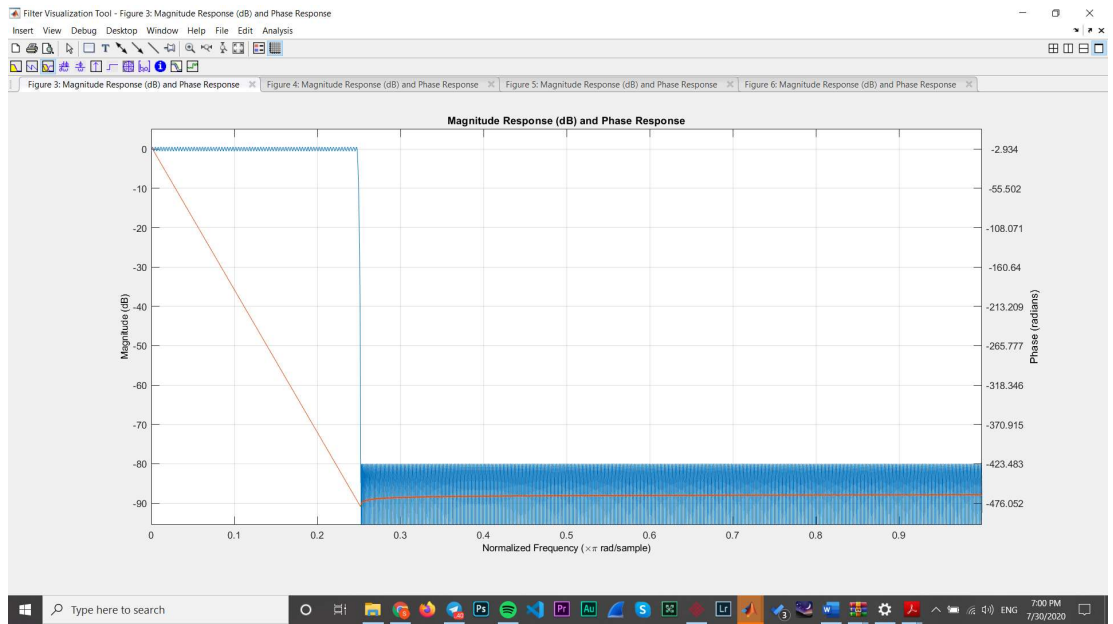
F_{stop} : 11950 , 18050

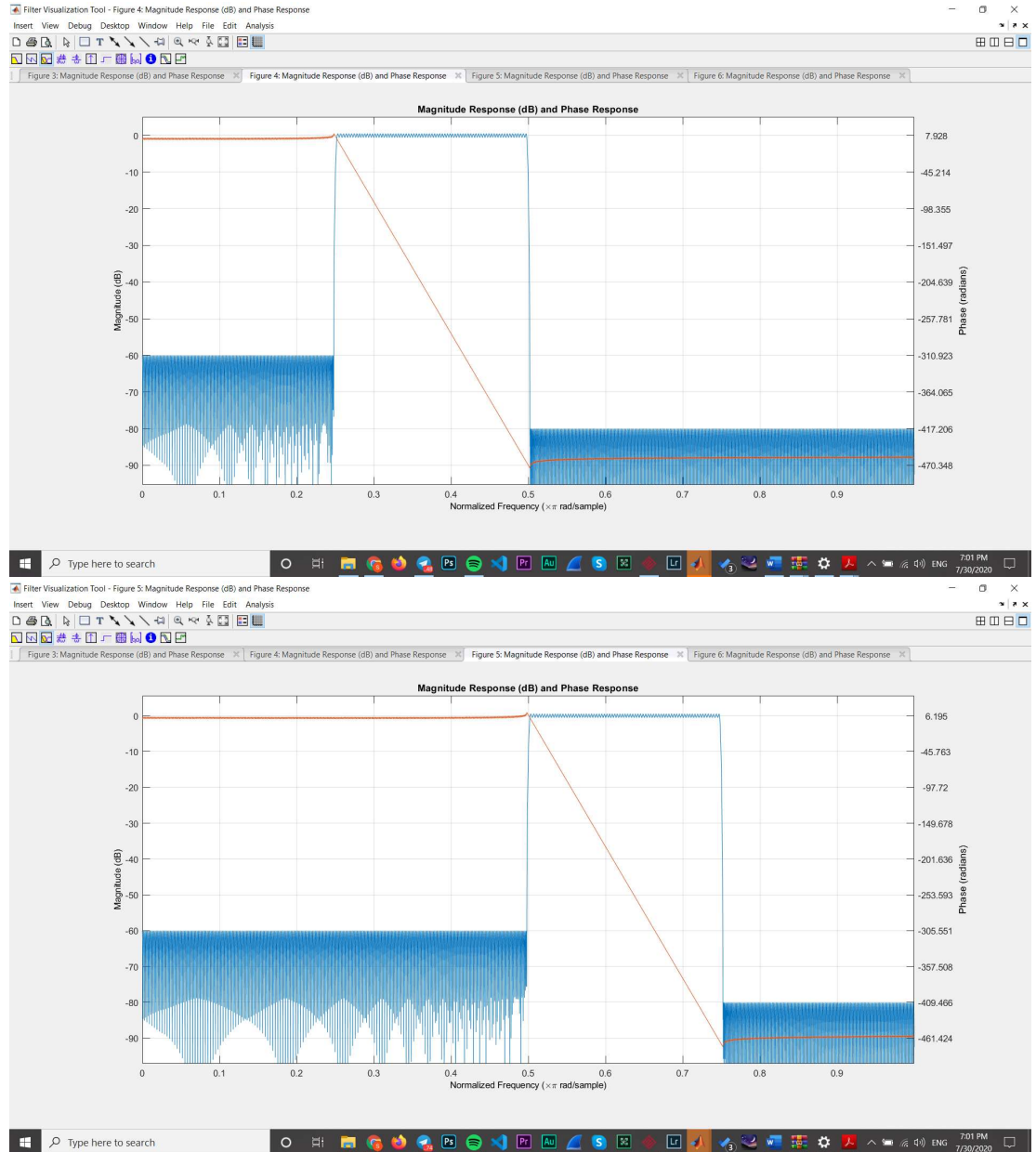
High pass filter :

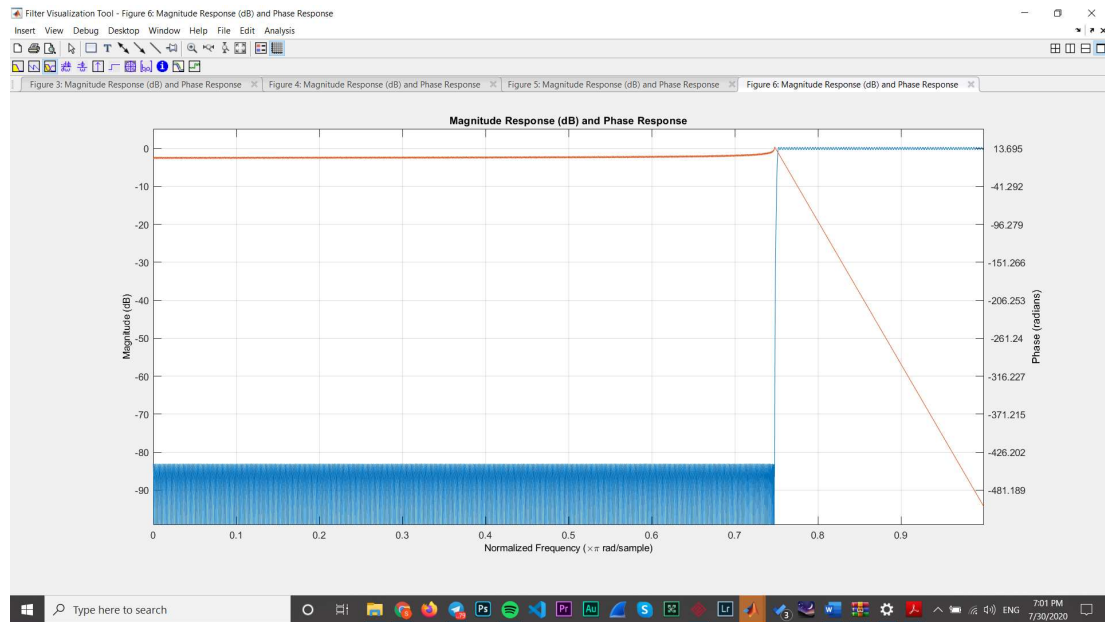
F_{pass} : 24050

F_{stop} : 23950

- نمودار اندازه و فاز پاسخ فرکانسی فیلتر های مذکور به صورت زیر است :

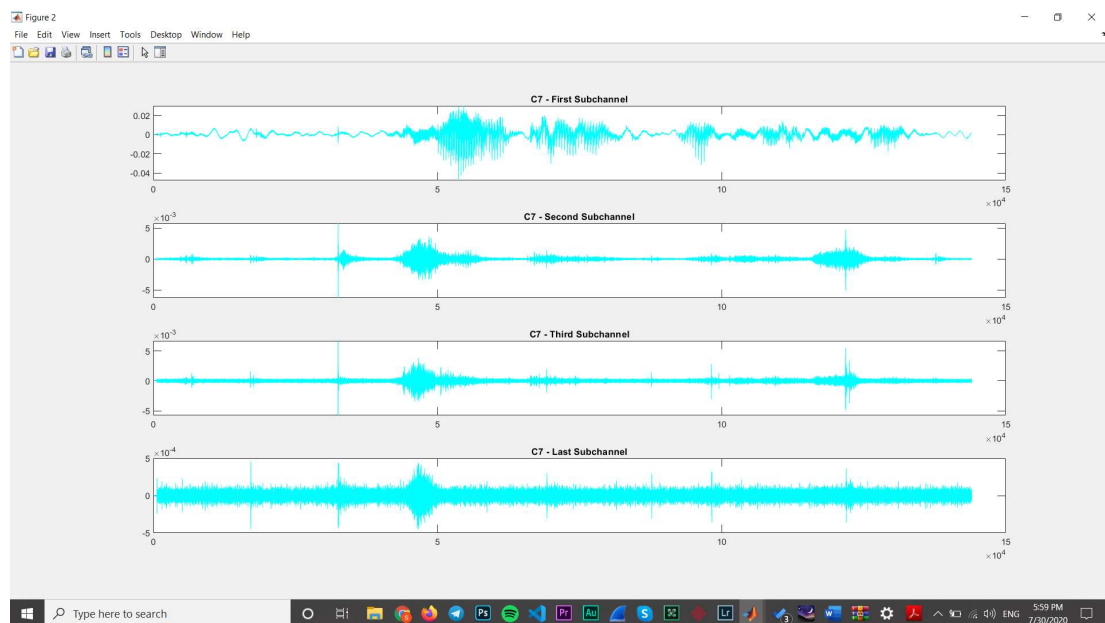


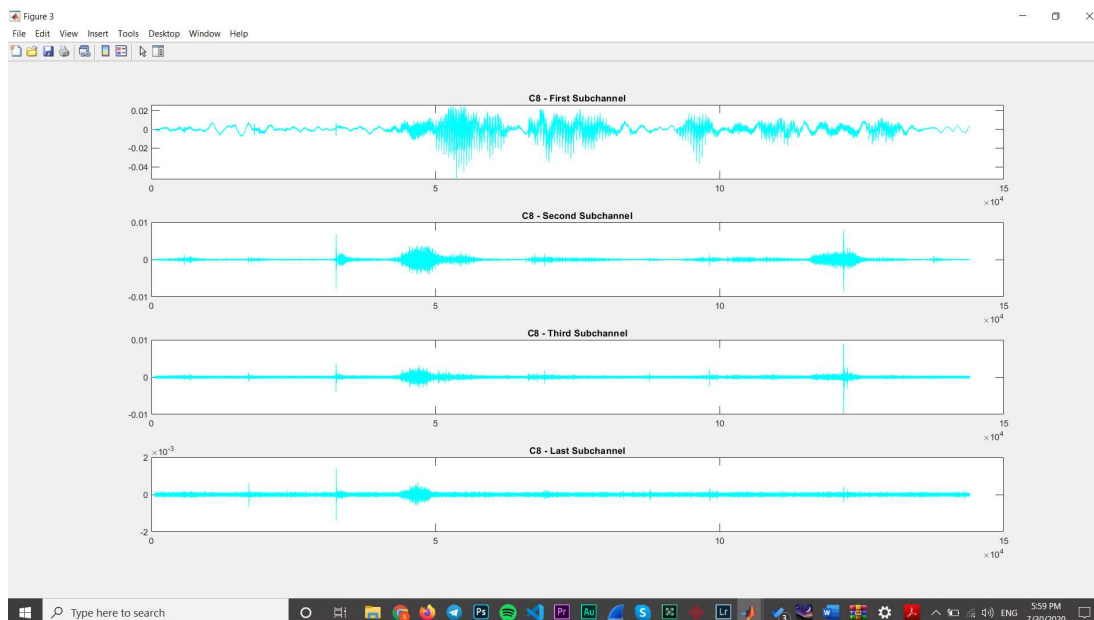




سوال ۳ –

همچنین طیف فرکانسی سیگنال های c7 , c8 پس از فیلتر شدن در subchannel های تولید شده توسط این ۴ بازه فرکانسی به شکل زیر است :



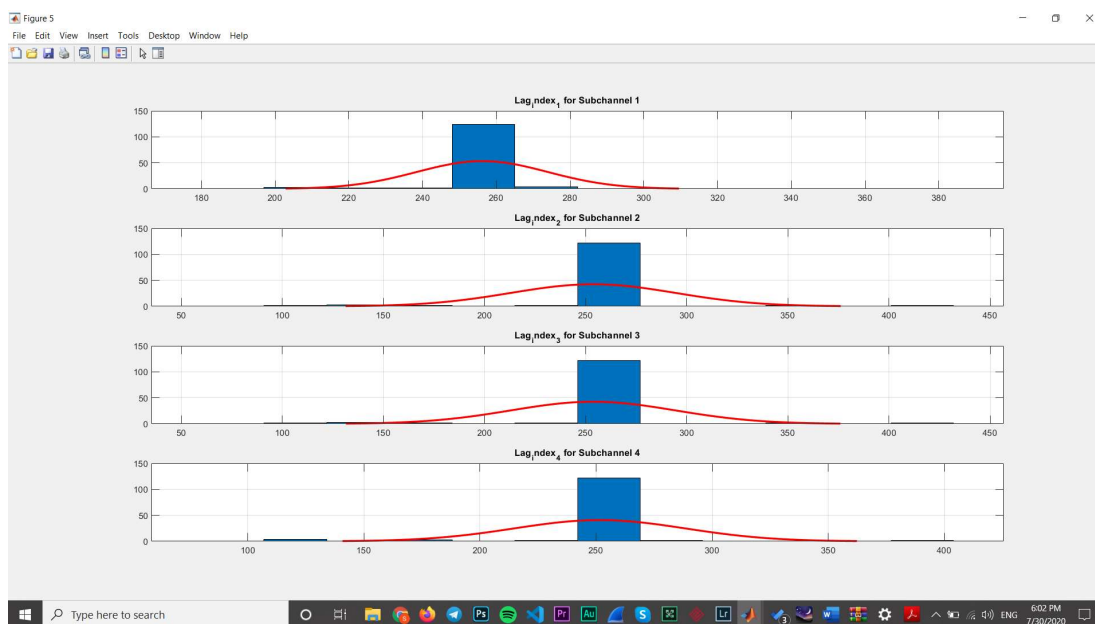
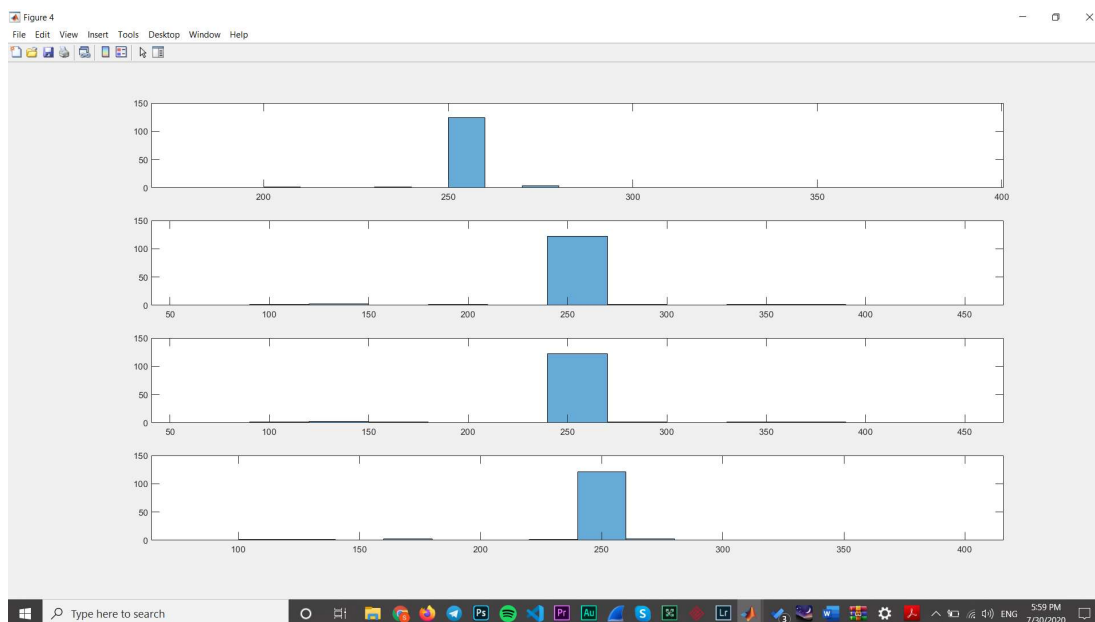


با توجه به اینکه محدوده فرکانسی صحبت انسان حدوداً در زیرکانال‌های اول و دوم است (تا 12kHz) مشاهده می‌شود که بیشتر انرژی سیگنال‌های فیلتر شده در این دو زیرکانال قرار دارد. سیگنال‌های هر کدام از زیرکانال‌ها به منظور کوتاه‌تر شدن حجم محاسبات با فاکتور 4 decimate می‌شوند. پردازش روی آن‌ها صورت می‌گیرد و در نهایت در قسمت نهایی با همین فاکتور بازیابی (و درونیابی) می‌شوند.

حال برای یافتن Lag index داده‌ها و حصول میانگین و واریانس آن‌ها ابتدا سیگنال‌های زیرکانال‌ها را مطابق دستور کار گفته شده به chunk‌های کوچکتر تقسیم می‌کنیم که در این مورد چون داده‌ها decimate شده‌اند و طول هر chunk 256 واحد است، 140 چانک برای هر کدام از زیرکانال‌ها خواهیم داشت که در یک حلقه for داده‌های زیرکانال‌ها به آن‌ها assign می‌شوند.

سوال ۴ و ۵ –

در ادامه بین چانک‌های متناظر در هر زیرکانال با هم correlation گرفته می‌شود و توزیع lag index از کنار هم قرار دادن حاصل correlation بین چانک‌ها در یک ماتریس و کشیدن هیستوگرام آن‌ها در بازه‌ای که داده‌های پرت چندانی وجود ندارد (xlim) بدست می‌آید.



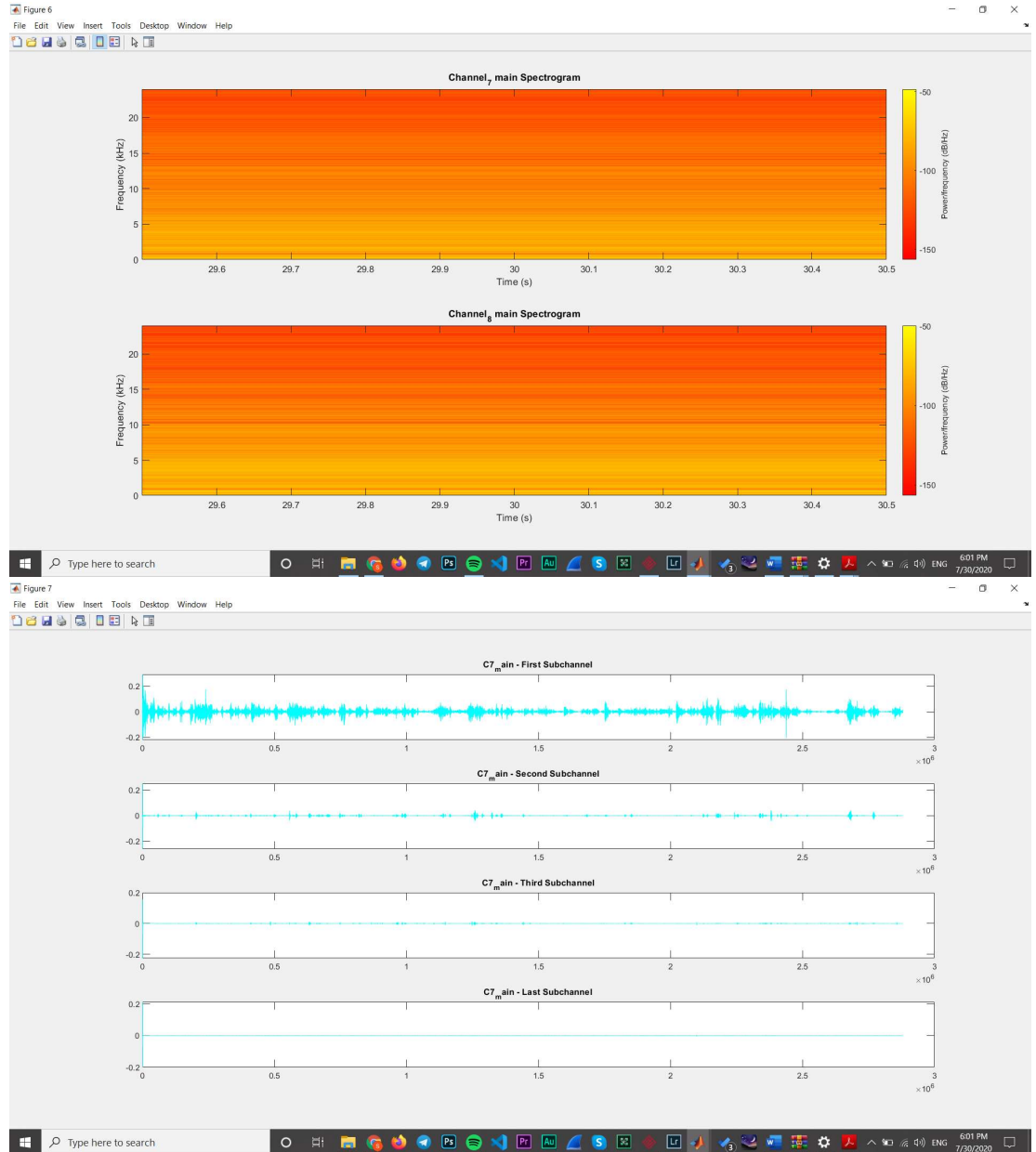
در صورتی که به هیستوگرام مربوط به lag index ها یک توزیع نرمال فیت کنیم (با استفاده از distfit) و میانگین آن را هم با تقریب در نظر بگیریم. به کمک روش یک سیگما (68.27 درصد داده های زیر نمودار از میانگین به دو طرف که توضیح آن به صورت مختصر در کامنت های فایل تابع deviation.m و فایل area.m که تابعی برای محاسبه مساحت کل هیستوگرام مربوط به lag index ها در بازه داده شده است)، انحراف معیار حدودی توزیع lag index در هر کدام از زیر کانال ها بدست می آید. که در شکل زیر مقادیر انحراف معیار های محاسبه شده برای هر کدام از زیر کانال های ۱ تا ۴ در فایل های intro و main مشاهده می شود.

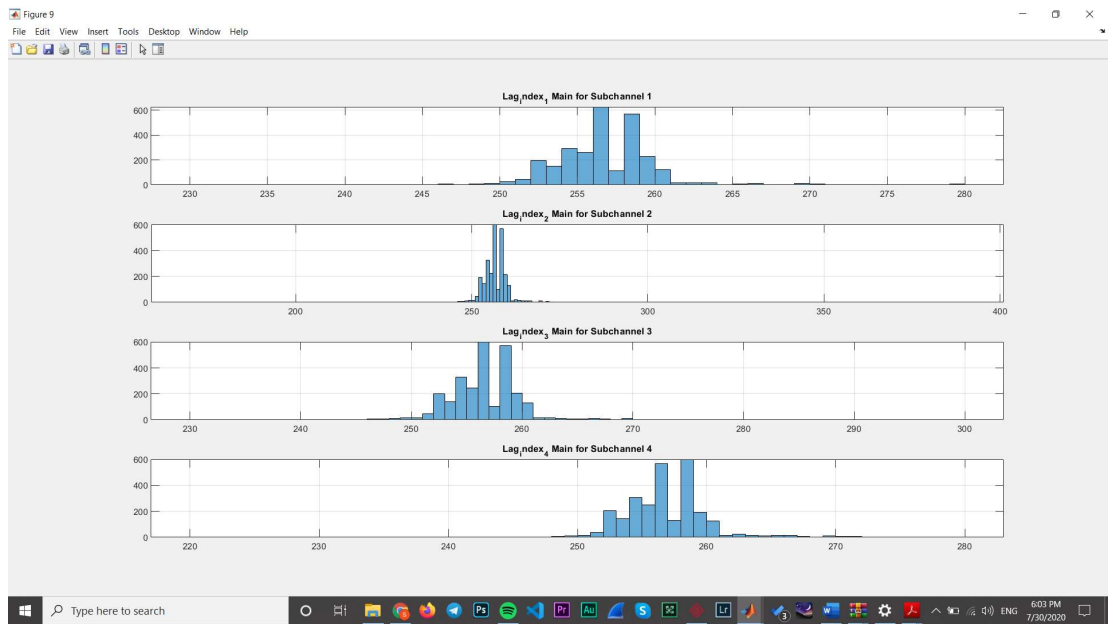
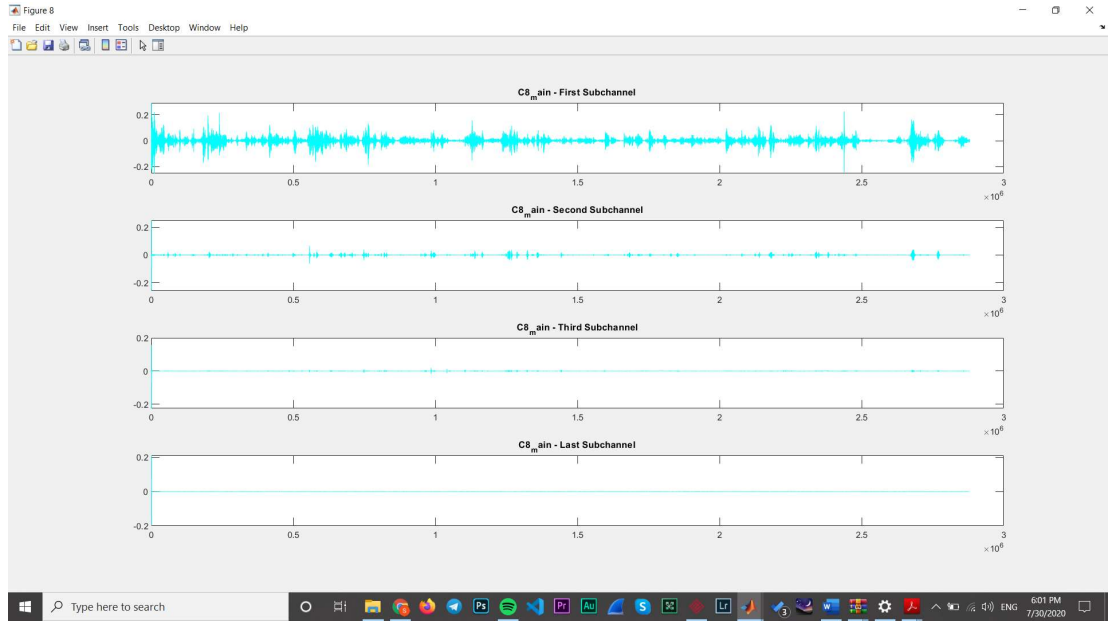
std_dev_1	0.8286
std_dev_1_m	3
std_dev_2	2.2429
std_dev_2_m	3
std_dev_3	0.8286
std_dev_3_m	3
std_dev_4	0.8286
std_dev_4_m	3
sub_ch_corr	140x4 double
sub_ch_corr_m	2812x4 double
total_area_1	127
total_area_1_m	2781
total_area_2	125
total_area_2_m	2801
total_area_3	123
total_area_3_m	2760
total_area_4	135
total_area_4_m	2756

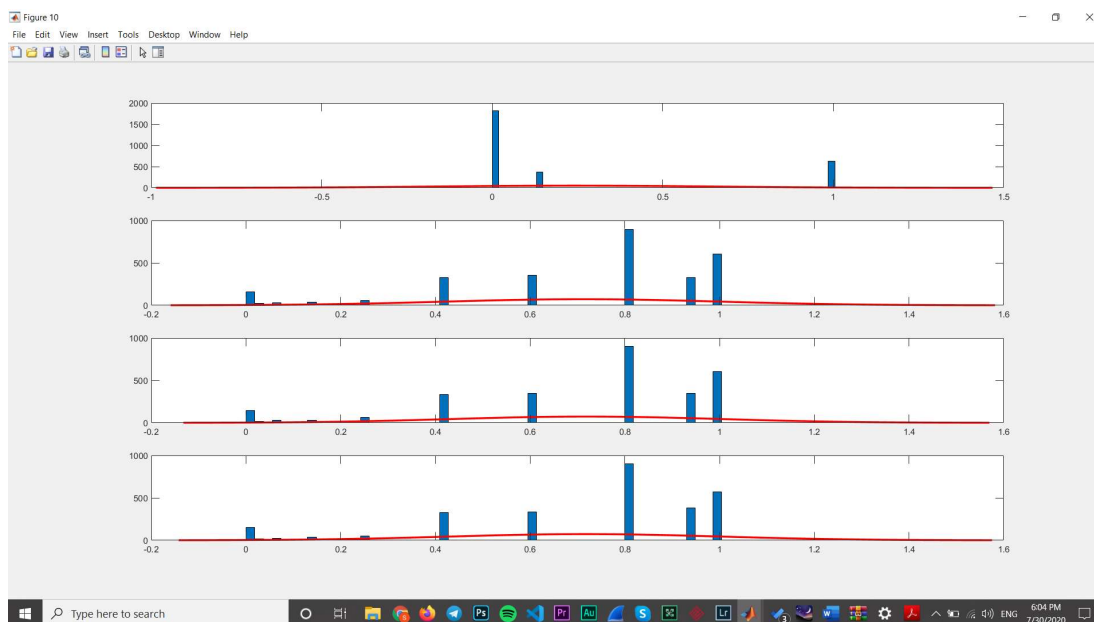
لازم به ذکر است برای آنکه صدای هدف (خروجی) جدا کردن صدای گوینده مرد از سایر صدا ها است وقتی که کانال های ۷ و ۸ روبروی گوینده هدف قرار دارند و تنها با اختلاف فاز کمی سیگنال صدای او به گیرنده ها می رسد در صورت کورلیشن گرفتن بیشترین تمرکز lag index برای صدای هدف حول میانگین توزیع خواهد بود، بنابراین به کمک کوچکتر گرفتن بازه رسم هیستوگرام و حذف داده های پرت که تاثیر بالایی روی محاسبه انحراف معیار دارند، با ضرایب W دقیقتری در ادامه مواجه خواهیم شد.

سوال ۶ –

در این سوال خواسته شده تا مراحل ۱ تا ۴ را برای سیگنال اصلی نیز تکرار کنیم. که از آنجا که برای محاسبه ی ضرایب w در سوال ۷ نیاز به داشتن واریانس و میانگین توزیع lag index می باشد قسمت ۵ را هم برای سیگنال main تکرار می کنیم.







سوال ۷ الی ۱۰ (بازیابی سیگنال)

در قسمت ۷ :

میانگین و واریانس بدست آمده در قسمت قبل را برای محاسبه ی ضرایب w استفاده میکنیم . با توجه به اینکه پس از decimation ، هر کدام از زیر کانال های سیگنال main حاوی 2812 چانک شد باید برای هر کدام از این چانک ها به طور جداگانه w محاسبه شود، ماتریس w عملاً یک ماتریس 3 بعدی است که بعد سوم شامل 8 طبقه برای هر یک از زیرکانال های $c7, c8$ ، هر طبقه شامل 2812 ستون که هر یک نشان دهنده چانک ها هستند و هر ستون در واقعاً نماینده ضریب مربوط به یک چانک می باشد، است.

(برای حصول نتیجه بهتر و در آوردن صدای هدف به صورت بهتر سیگمای زیر کانال ۱ حدود ۰.۵/ گرفته شد) به صورت دستی و به روش آزمون و خطا))

در قسمت ۸ :

از سیگنال main که پردازش روی آن انجام شده و وزن ها روی سیگنال اعمال می شود را خروجی می گیریم و با توجه به گسستگی صدا تکه تکه به نظر می رسد و این مشکل را با کمک فیلتر smooth در قسمت بعد تصحیح می کنیم.

در قسمت ۹ :

با استفاده از فیلتر آماده smooth که از میانگین گیری داده های همسایه استفاده می کند تا کورلیشن (محتوای فرکانسی) را حول فرکانس های کوچک تقویت کند (به عنوان نوعی فیلتر پایین گذر عمل می کند) خدشه حاصل از گسستگی صدا و تغییرات ناگهانی سطوح صدا را کاهش می دهیم.

در قسمت ۱۰ :

باید بازیابی نهایی سیگنال را برای کانال های راست و چپ (۷ و ۸) انجام دهیم که چون decimation انجام شده بود باید interpolation را در اینجا با فاکتور ۴ انجام دهیم تا طول سیگنال ثابت بماند. هر چند در این حالت عملاً یک چهارم داده های خام پردازش می شوند و از درونبایی آن های سایر داده ها که تشکیل دهنده سیگنال خروجی اند بدست می آید.