

بسم تعالی



پردازش سیگنال های الکتروانسفالوگرام

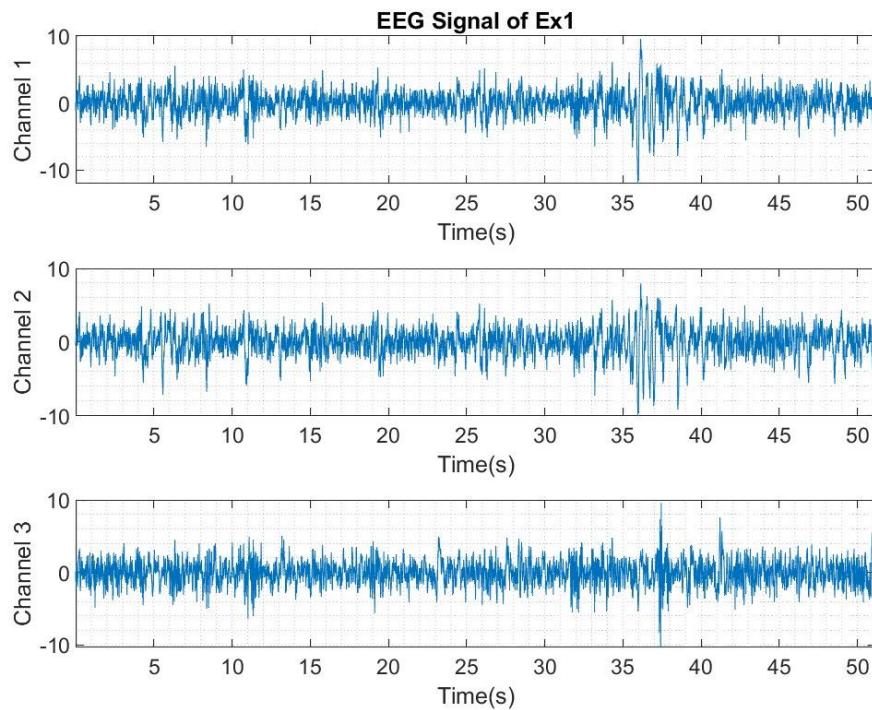
تمرين کامپیوتري ۲

امیرحسین زاهدی ۹۹۱۰۱۷۰۵

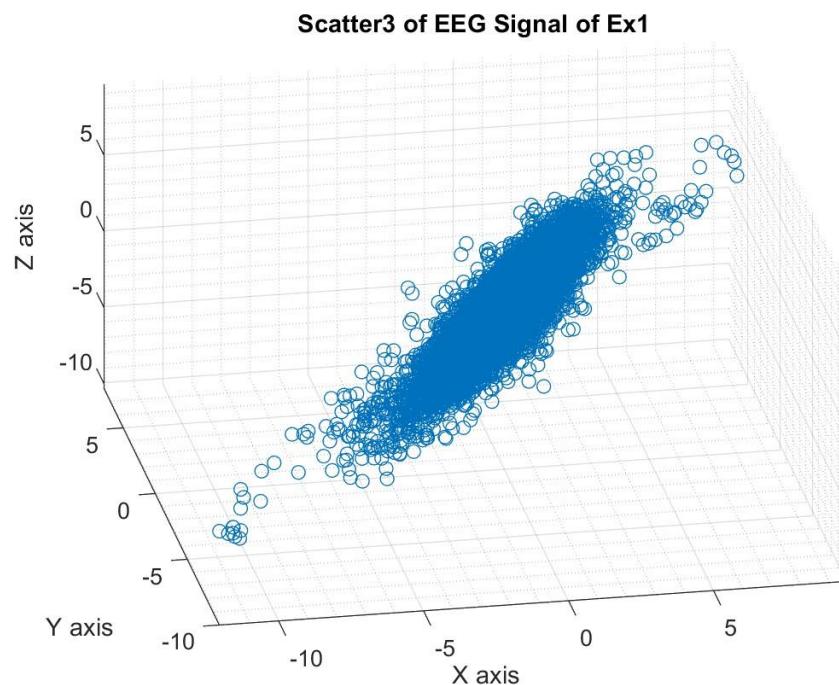
پاییز ۱۴۰۲

سوال ۱:

الف) سیگنال را در سه کanal رسم می کنیم.



ب) با استفاده از دستور scatter3 داده ها را در سه بعد بر حسب کانال های ۱ تا ۳ رسم می کنیم.



X و Y و Z فوئدی از کanal های ۱ تا ۳ هستند. داده ها در هر سه جهت پخش هستند اما در صفحه x-y گویا که یک رابطه خطی دارند و می توان آن ها را با یک خط مدل کرد. احتمالا در PCA، خط گفته شده به عنوان کامپوننت اول بدست می آید. و کامپوننت دوم نیز در جهت z و عمود بر کامپوننت اول باشد. کشیدگی در جهت x و y نسبت به z مشهود تر است.

ج) ماتریس داده هایمان شامل سه کانال در 10240 نمونه با فرکانس نمونه برداری 200 هرتز است. پس به عبارتی در ماتریس مشاهدات سه کانال داریم که هر کدام یک بردار برای ماتریس هستند. با استفاده از حذف میانگین، و سپس بدست آوردن بردارهای ویژه و مقادیر ویژه از این سه بردار، می خواهیم که ماتریس را سفید سازی کنیم، به عبارتی می خواهیم که ترکیب خطی ای را بر روی ماتریس مشاهدات اعمال کنیم تا ماتریس کوواریانس خروجی برابر ماتریس همانی بشود. برای همین، ابتدا ماتریس ترانهاده بردارهای ویژه ماتریس کوواریانس مشاهدات را بدست آورده و در خودش ضرب می کنیم. سپس ماتریس مقادیر ویژه را با توان 0.5 - در ماتریس حاصل مرحله قبل ضرب می کنیم. اینگونه ماتریس کوواریانس خروجی برابر همانی می شود. پس اگر همان دو ماتریس گفته شده را در ماتریس مشاهدات ضرب کنیم، خروجی ای بدست می آوریم که کوواریانس همانی است. اینگونه PCA اعمال می شود. شاهد هستیم که بردارهای بدست آمده بر هم عمود هستند و همبستگی ندارند. همچنین بردارها به ترتیب از واریانس بیشتر به واریانس کمتر می روند که میزان پراکندگی داده ها در آن بردار، بیانی از انرژی در راستای آن بردار است. به این ترتیب بردارها به ترتیب از انرژی زیاد به کم مرتب می شوند.

$$\mathbf{y} = \mathbf{Bz} = \mathbf{BAx} = \mathbf{Dx} \rightarrow \mathbf{C}_y = \mathbf{BC}_z\mathbf{B}^T = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix}$$

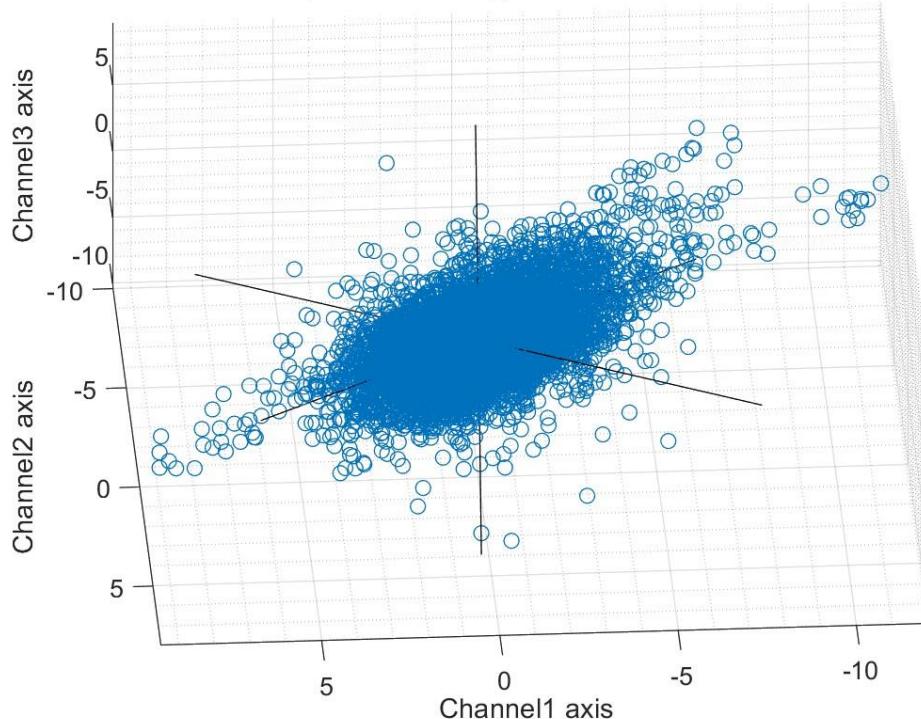
$$\mathbf{D} = \mathbf{BA} = \Lambda^{-2}\mathbf{U}'$$

$$\mathbf{C}_y = \mathbf{DC}_x\mathbf{D}^T = \Lambda^{-\frac{1}{2}}\mathbf{U}'\mathbf{C}_x\mathbf{U}\Lambda^{-\frac{1}{2}} = \mathbf{I}$$

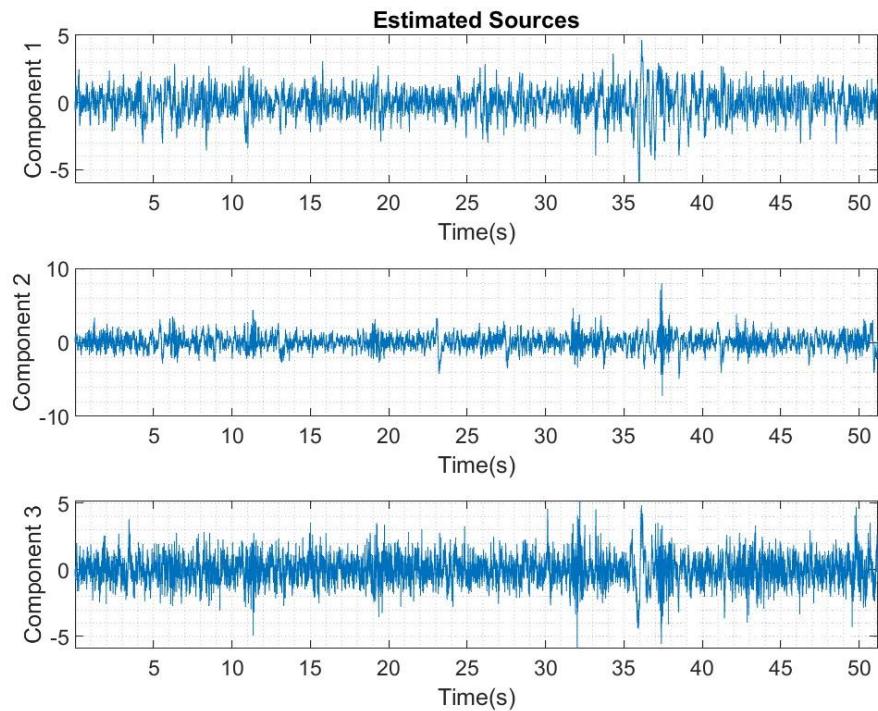
الگوریتم بالا را بر روی دیتای مشاهدات پیاده می کنیم تا با استفاده از PCA سه بردار عمود بر هم سفید سازی شده را بدست بیاوریم. این بردارها می توانند براورده از سورس های سازنده مشاهدات باشند.

پس از بدست آوردن بردارها، آن ها را بر روی سیگنال اصلی نمایش می دهیم:

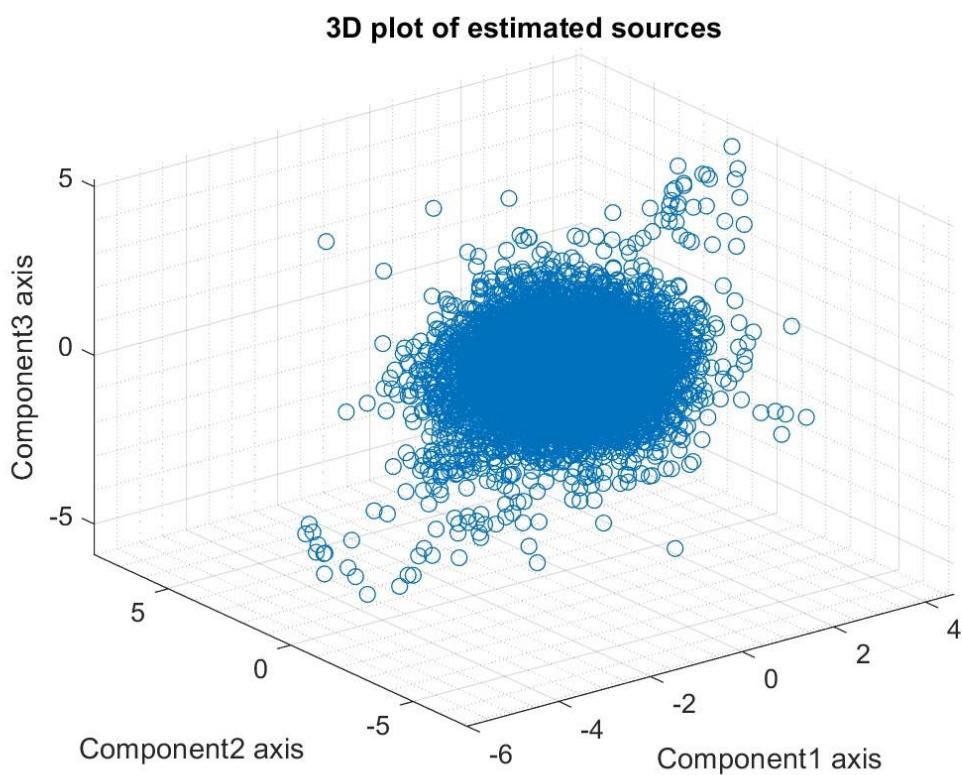
3D plot of EEG Signal and PCA vectors



حال هر کامپوننت را در زمان رسم می کنیم:



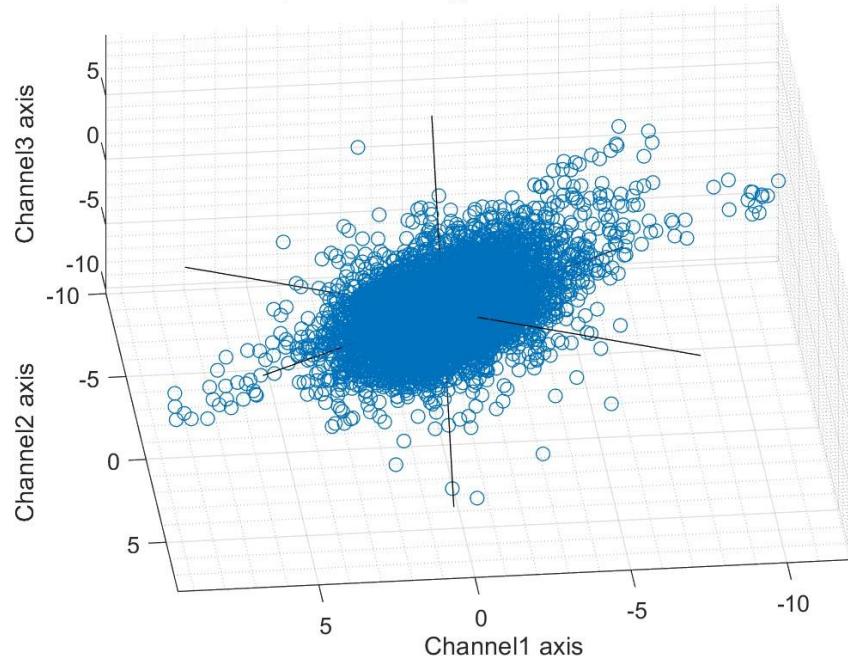
و در نهایت داده ها را در صفحه مختصات جدید رسم می کنیم:



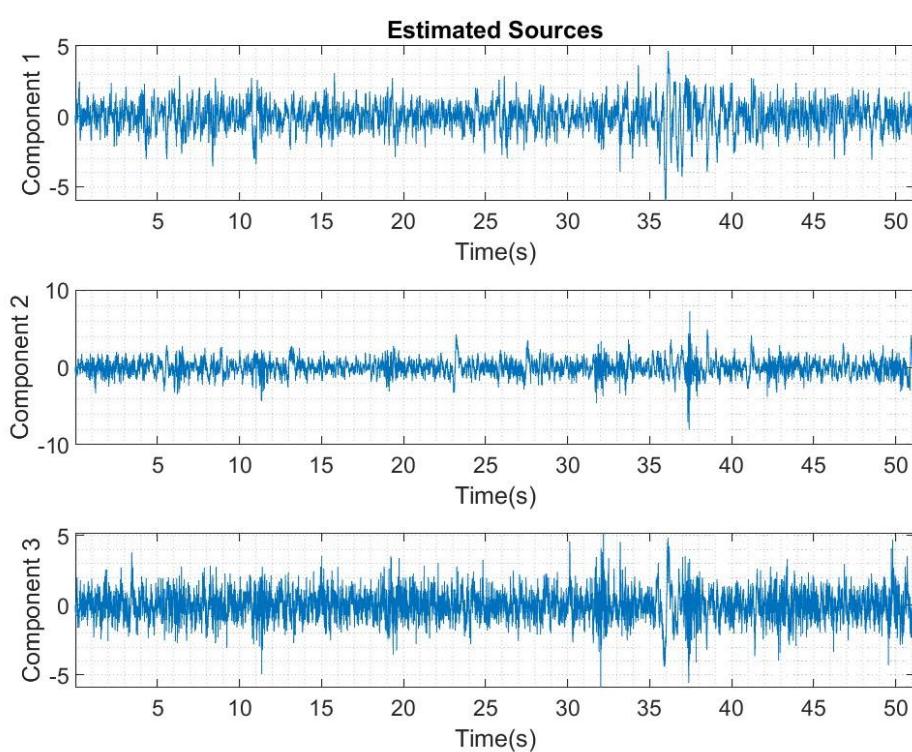
د) اینبار با استفاده از فانکشن پی سی ای خود متلب مراحل بالا را انجام می دهیم، ماتریس بردارها و مقادیر ویژه را بدست می آوریم. سپس ماتریس سورس ها را با استفاده از دیتای مشاهده شده بدست می آوریم.

پس از بدست آوردن بردارها، آن ها را بر روی سیگنال اصلی نمایش می دهیم:

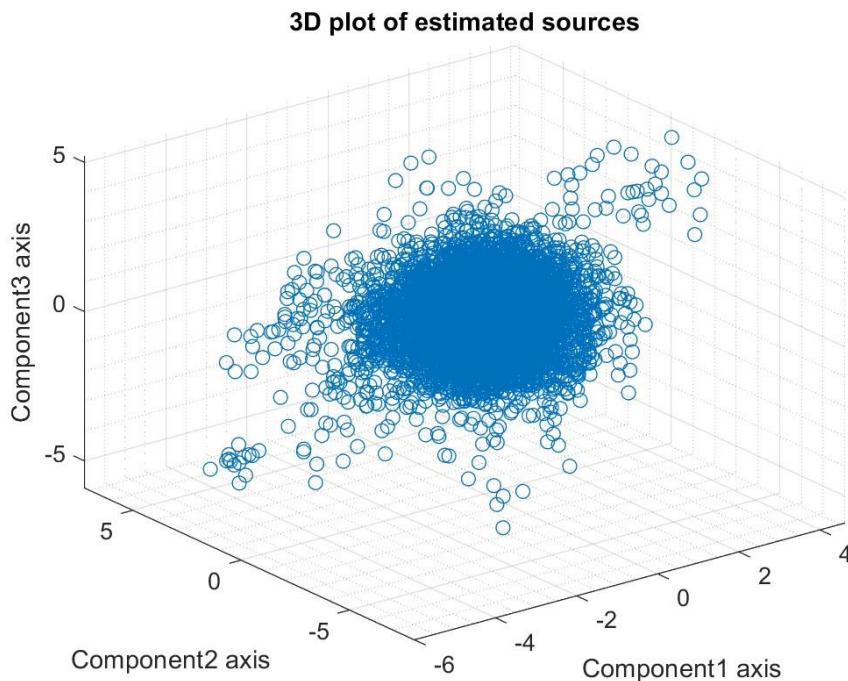
3D plot of EEG Signal and PCA vectors



حال هر کامپوننت را در زمان رسم می کنیم:



و در نهایت داده ها را در صفحه مختصات جدید رسم می کنیم:



۵) به صورت تئوری با اجرا کردن SVD بر روی سیگنال، می توانیم با استفاده از ماتریس های تکین راست و چپ و مقادیر تکین، برآورده از مقادیر ویژه و بردارهای ویژه ماتریس کوواریانس داشته باشیم. در PCA نیز داشتیم که بردارها و مقادیر ویژه ماتریس کوواریانس بست می آمدند و با استفاده از ترکیب مشاهدات و ماتریس بدست آمده سورس ها را تولید می کردیم. حال با اعمال کردن SVD و پیدا کردن مقادیر تکین می توانیم برآورده از لاندا که ماتریس قطری مقادیر ویژه بود داشته باشیم. همچنین، ماتریس تکین چپ (اگر ماتریس داده های 3×10240 را به عنوان ورودی بدهیم). همان ماتریس بردارهای ویژه کوواریانس مشاهدات است. پس می توان کشیدگی و واریانس داده ها و داده های سفید سازی شده را با روش SVD بدست آورد.

به عبارت ریاضیاتی، اگر ماتریس تکین چپ بدست آمده U ، ماتریس مقادیر تکین بدست آمده S ، ماتریس مقادیر ویژه کوواریانس مشاهدات Λ و ماتریس بردارهای ویژه کوواریانس مشاهدات C باشد. موقع داریم که:

$$C = U \Lambda U^T \quad \text{و} \quad \Lambda = (S^2)/(N-1)$$

با توجه به خروجی متلب، میبینیم که U و C برابر هم اما قرینه اند. و $\Lambda = \text{diag}(S^2)$ است.

```
U =
-0.6585  0.0299 -0.7520
-0.6385  0.5066  0.5794
-0.3982 -0.8617  0.3145
```

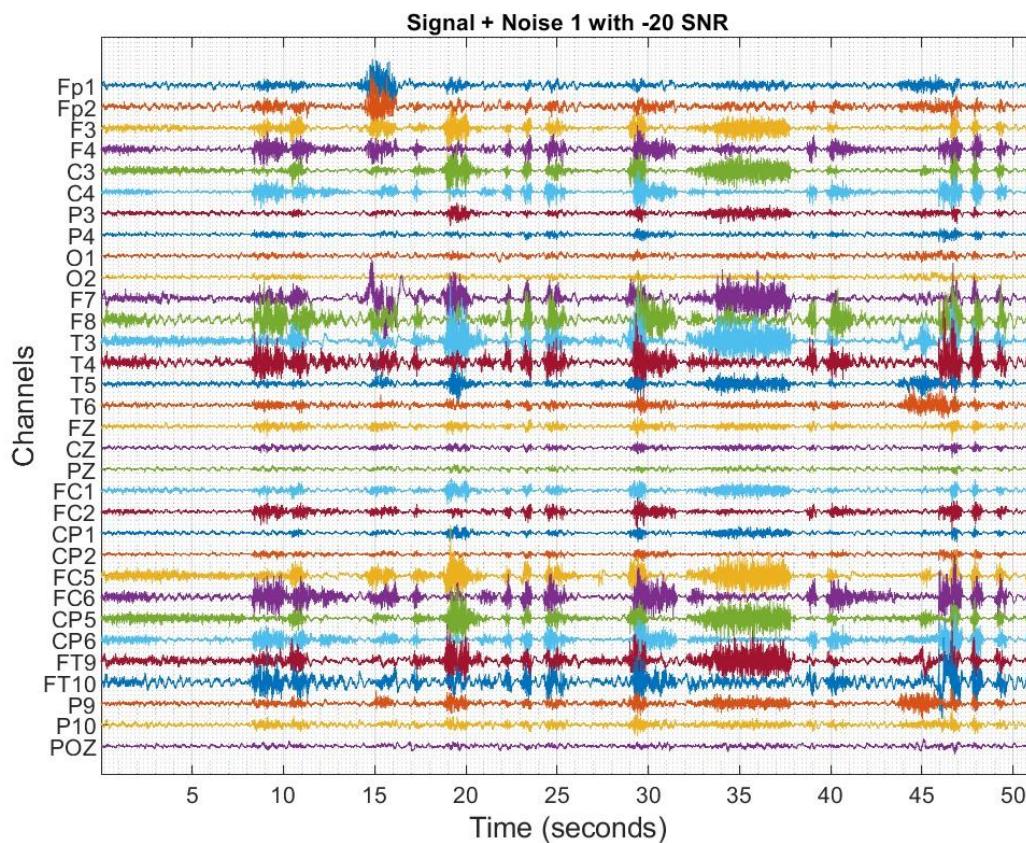
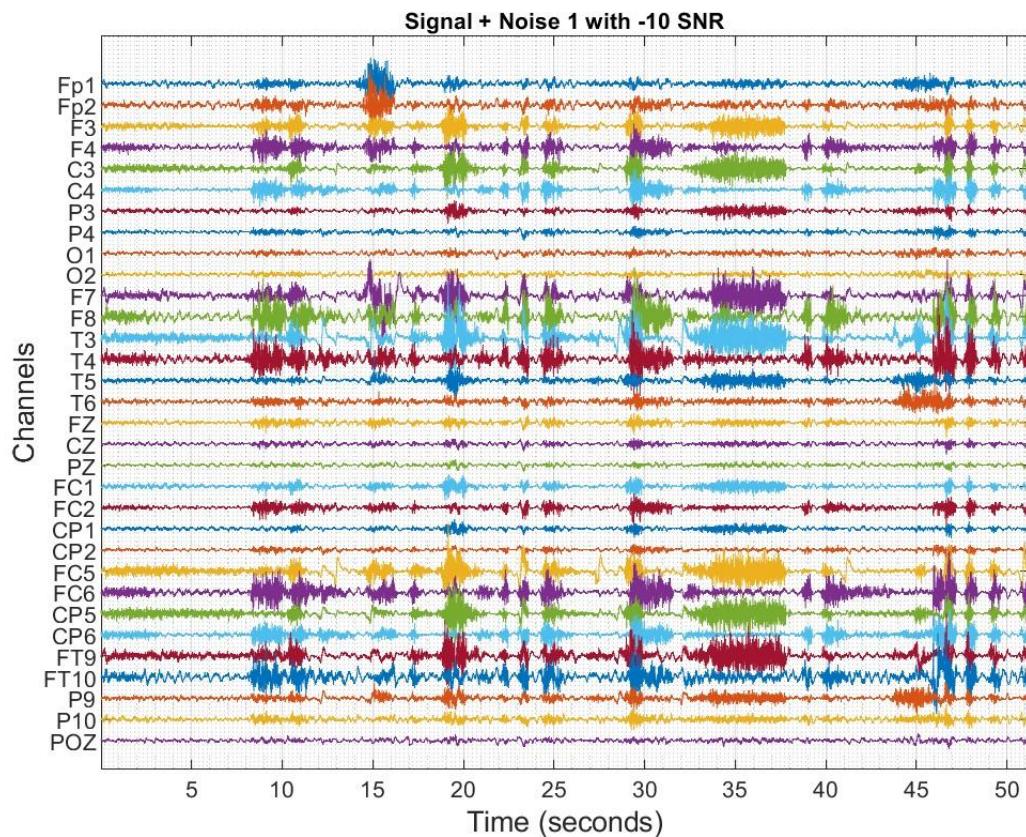
```
C =
0.6585 -0.0299  0.7520
0.6385 -0.5066 -0.5794
0.3982  0.8617 -0.3145
```

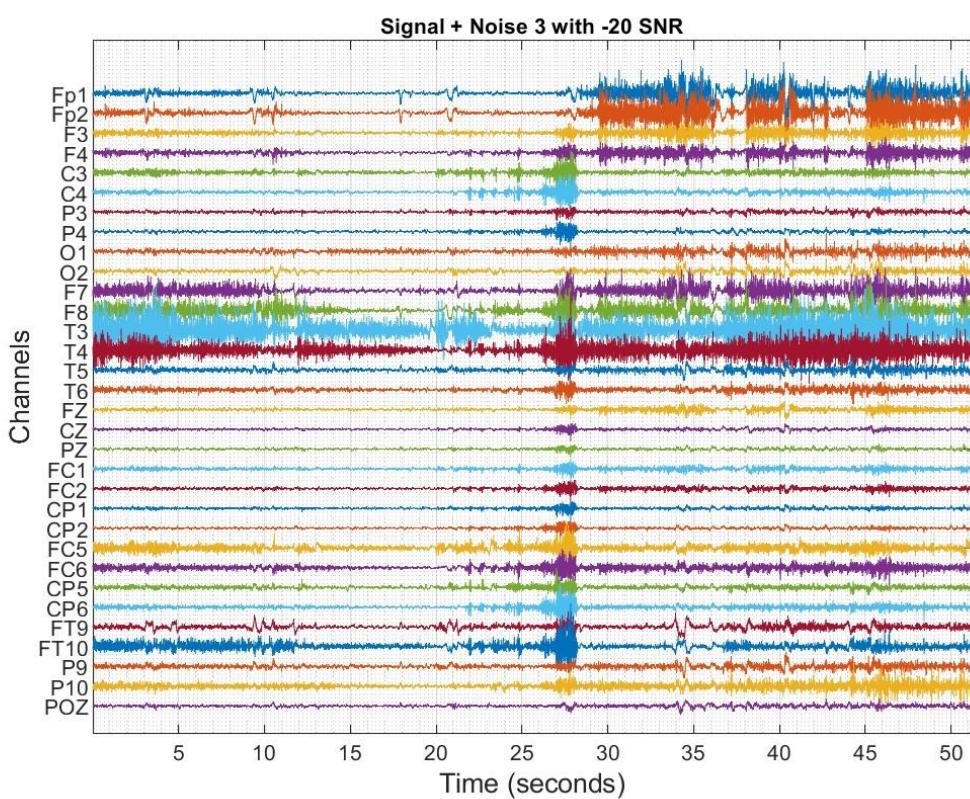
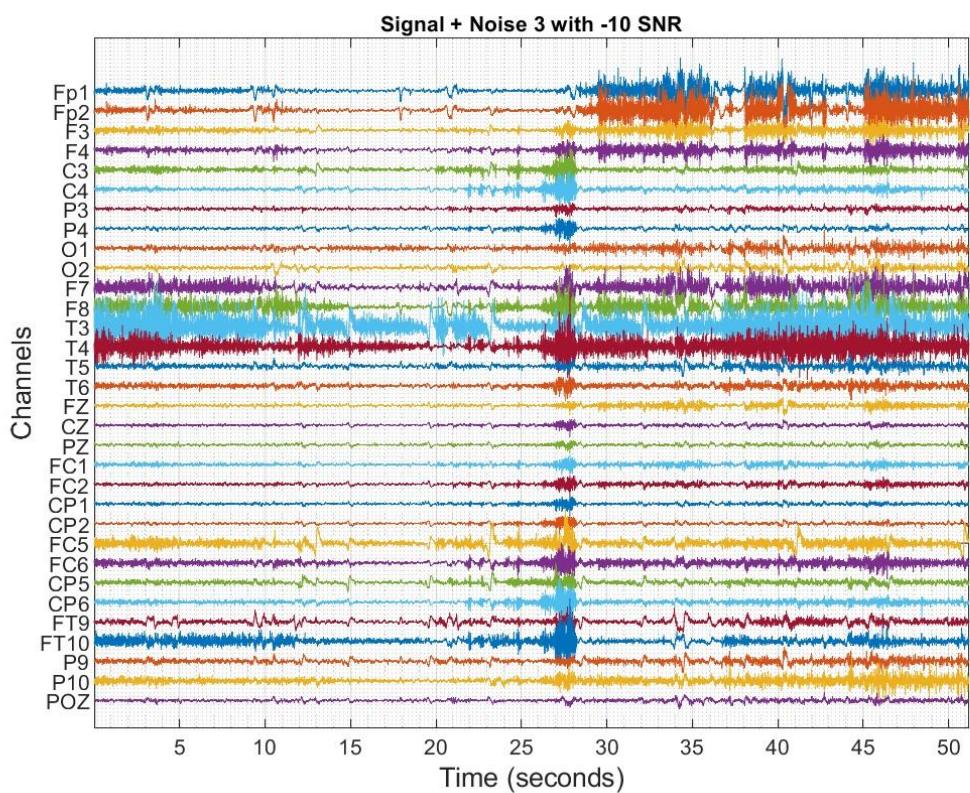
```
Lambda =
7.0891      0      0
0     1.7321      0
0      0     0.3333
```

```
S^2 / (N-1)
ans =
7.0891
1.7321
0.3333
```

سوال ۲:

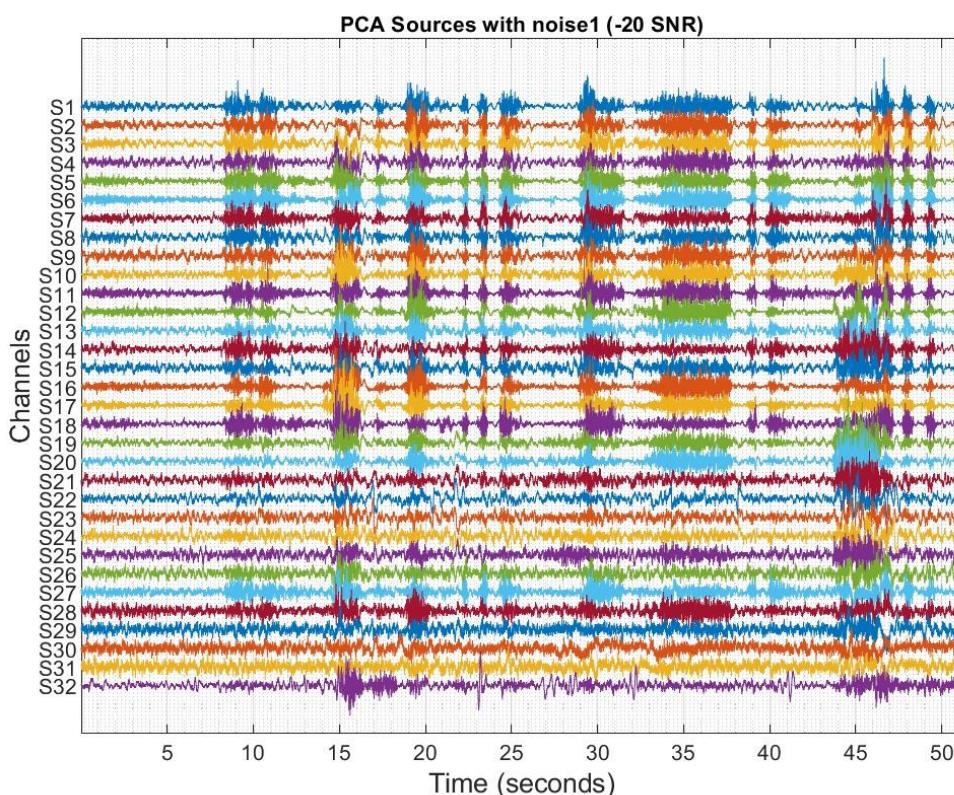
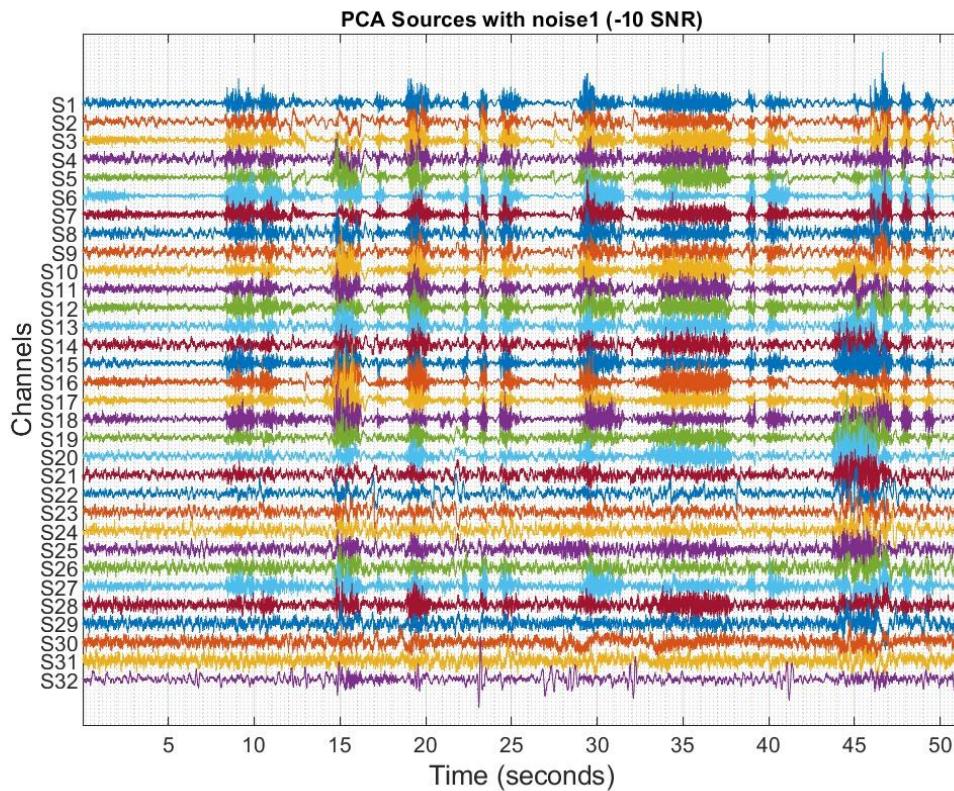
الف) نویز های ۱ و ۳ را با SNR های ۱۰- و ۲۰- با دیتای اصلی ترکیب کرد و سپس رسم می کنیم:

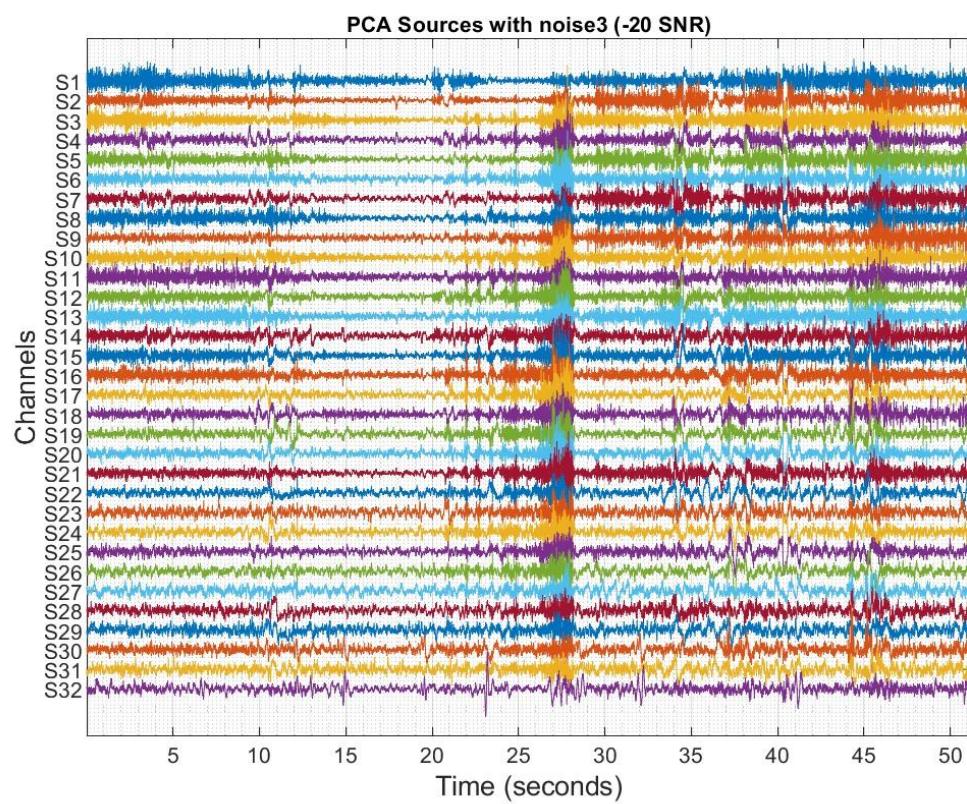
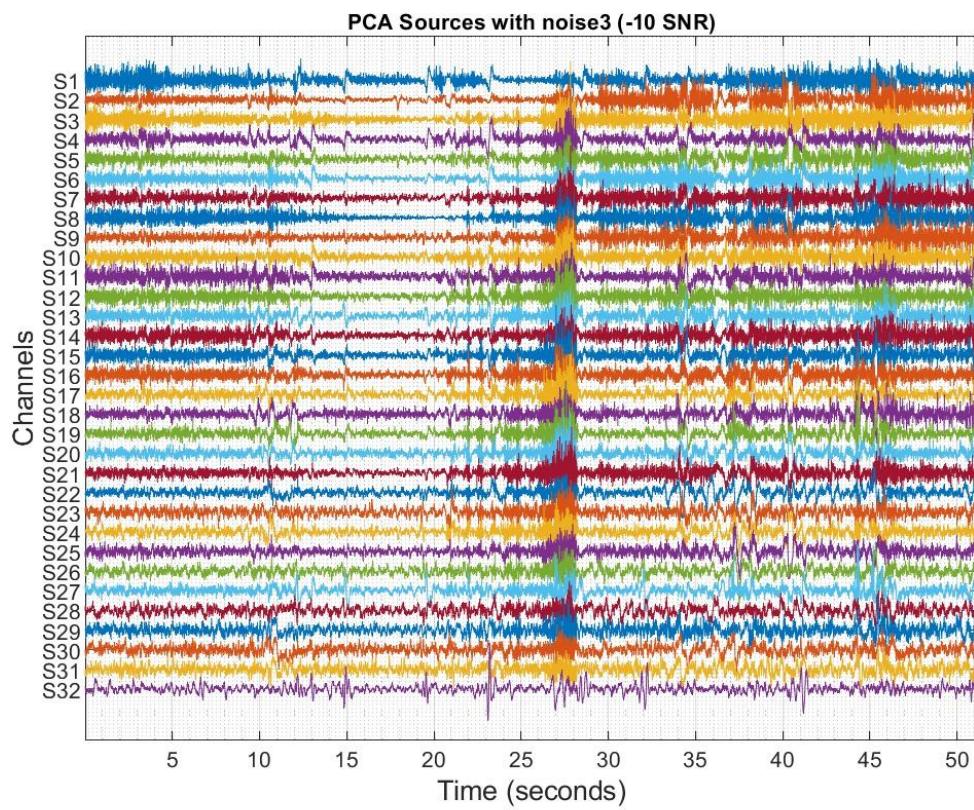




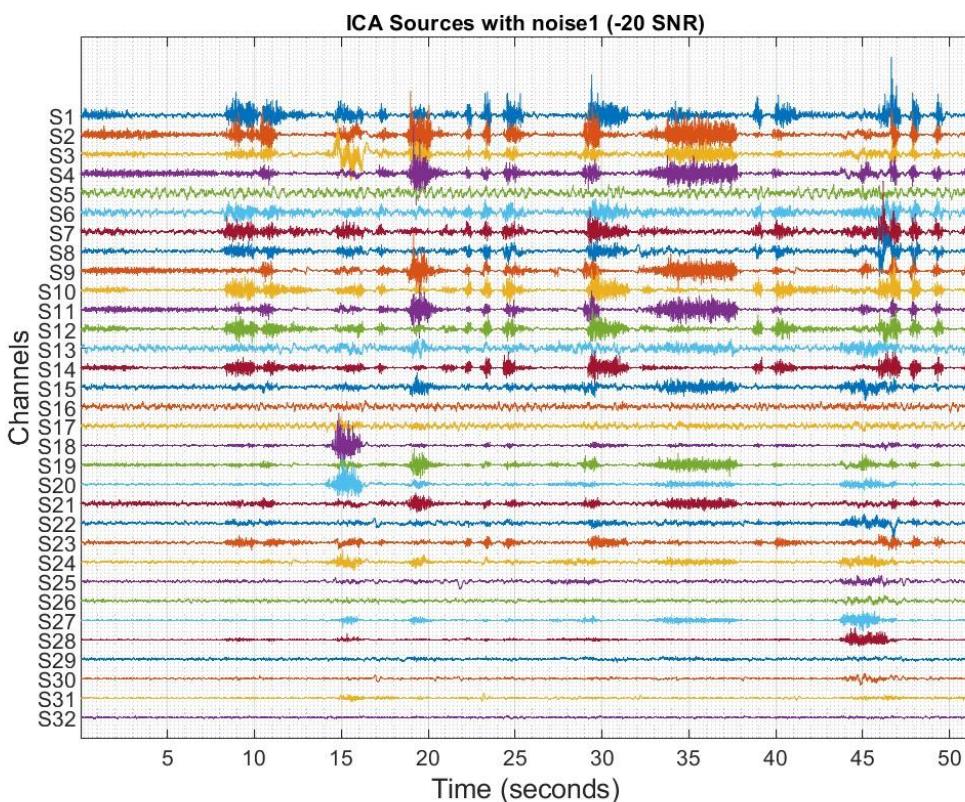
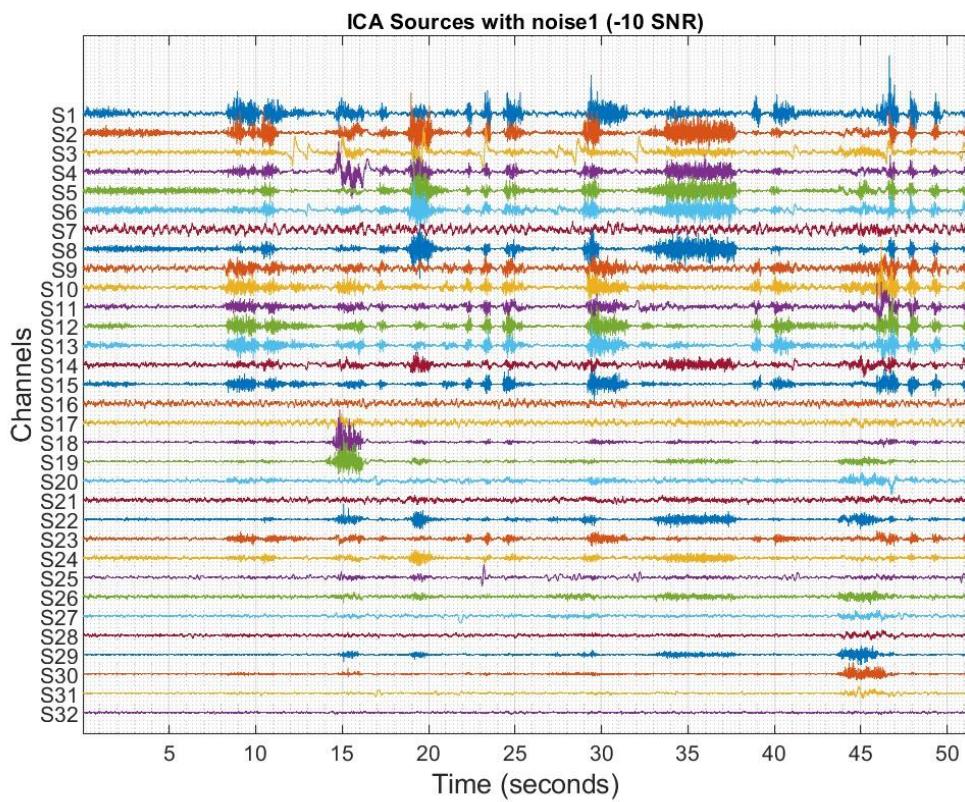
تفاوت دو نویز متفاوت که بسیار مشهود است اما در اس ان آر های مختلف گویا چندان فرقی دیده نمی شود و شکل سیگنال ها شبیه است. البته احتمالی وجود دارد که نویز بسیار غالب باشد و به دلیل غلبه نویز بر بک گراند ای جی، به دلیل به نمایش در نیامدن دامنه ها، تفاوت بین اس ان ارها دیده نمی شوند. با این حال $\text{SNR} = -20$ نسبت به $\text{SNR} = -10$ ، میزان تاثیر گذاری نویز ۱۰ برابر دارد.

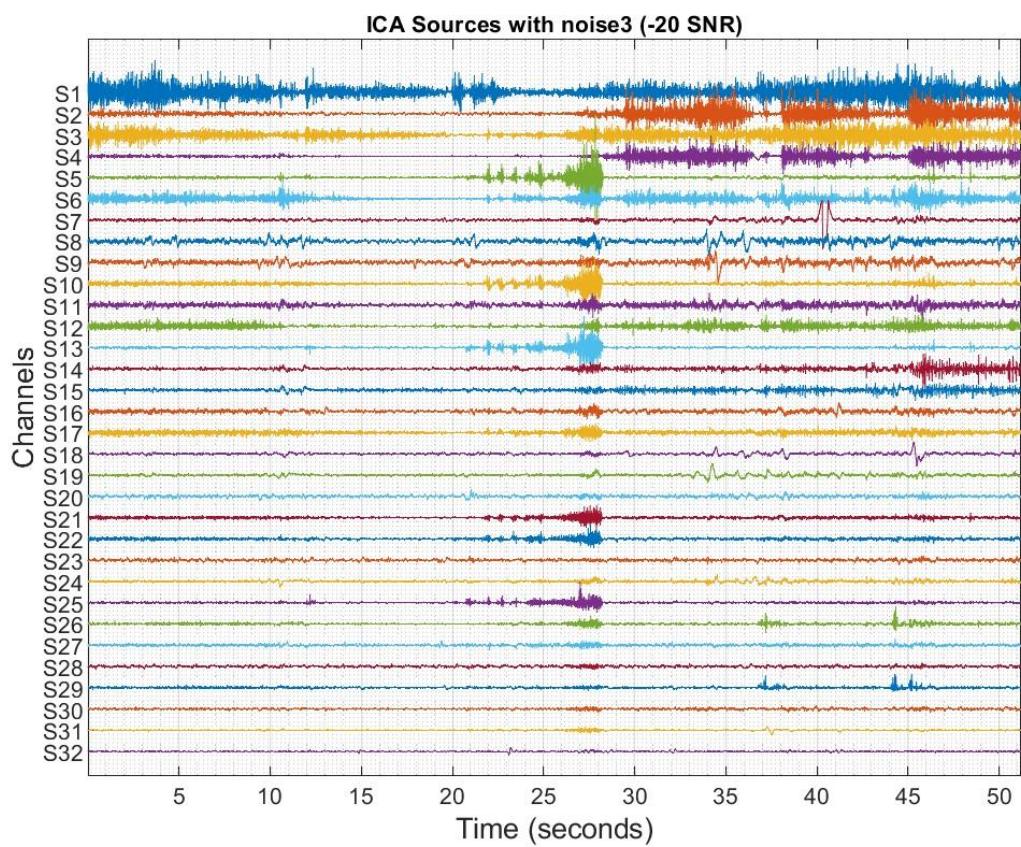
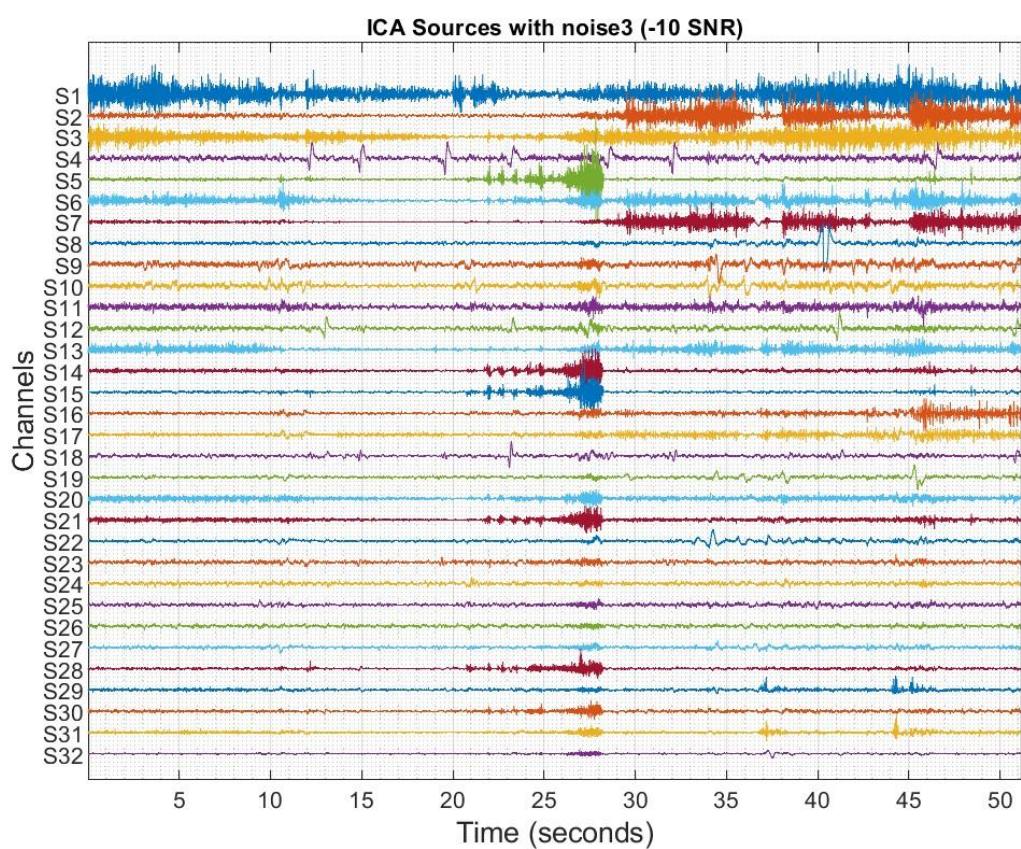
ب) با استفاده ازتابع pca که در سوال ۱ نیز استفاده شد، بردارهای ویژه و مقادیر ویژه ماتریس کوواریانس سیگنال EEG را بدست آورده و با استفاده از آن ها تخمینی از سурс های سیگنال های مشاهدات همراه با نویز بدست می آوریم که به شکل زیر هستند.





همان مراحل قبل را برای ICA با استفاده از تابع COM2R انجام می دهیم. سورس های بدست آمده به صورت زیر هستند.



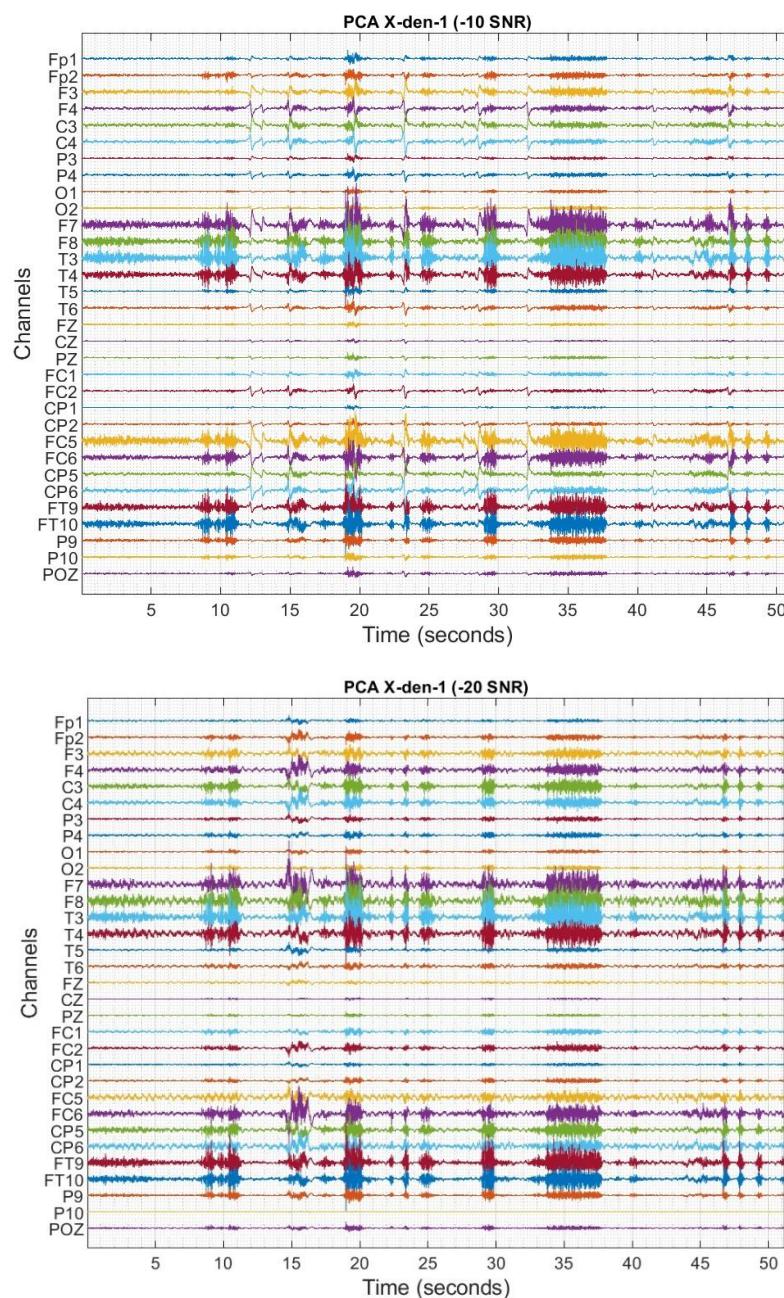


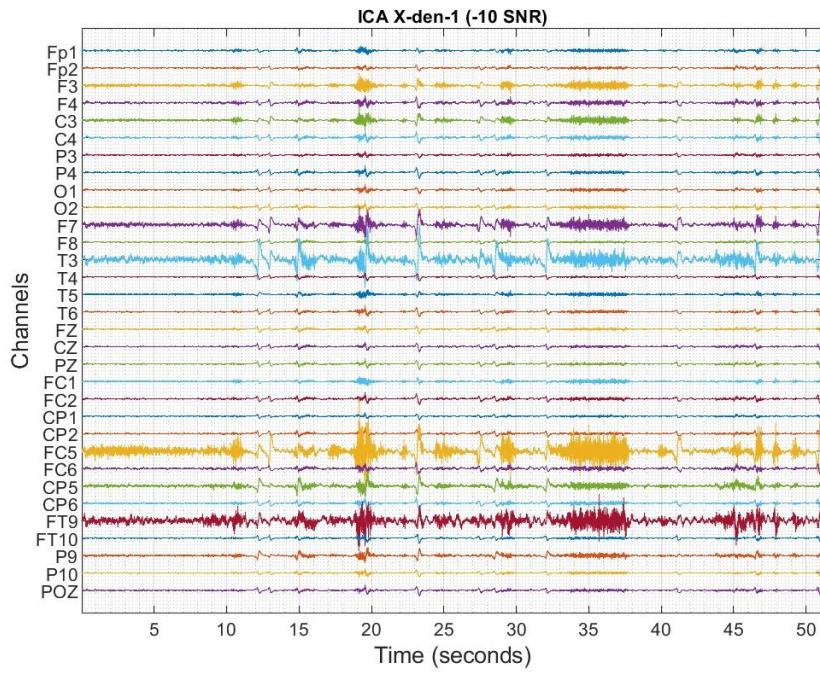
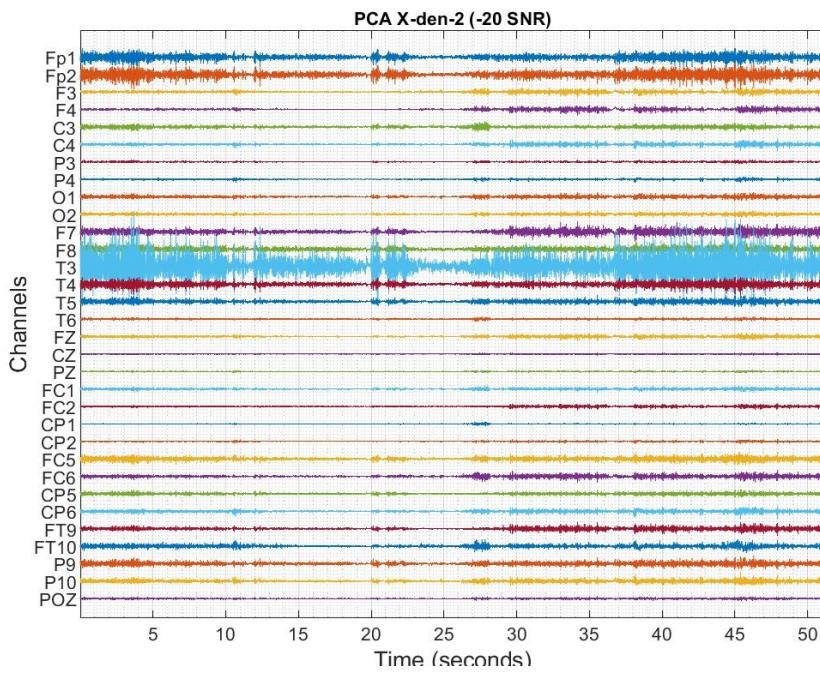
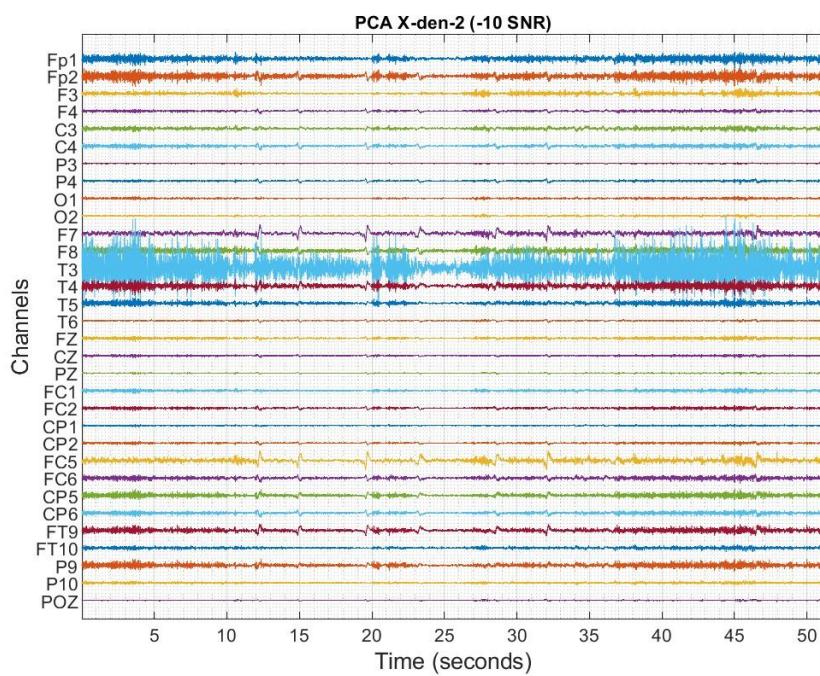
اگر بخواهیم دو روش را با یکدیگر مشاهده کنیم میبینیم که در PCA سورس ها نتیجه قابل قبولی نیستند و بسیار نویزی هستند اما در ICA این اتفاق بسیار کمتر مشاهده می شود. با وجود نویز بسیار کمتر در روش دوم می توان تا حدی منابع مطلوب و نا مطلوب را از یکدیگر تمیز داد اما در روش اول این اتفاق ممکن نیست و یا با خطای بسیار زیادی همراه خواهد بود.

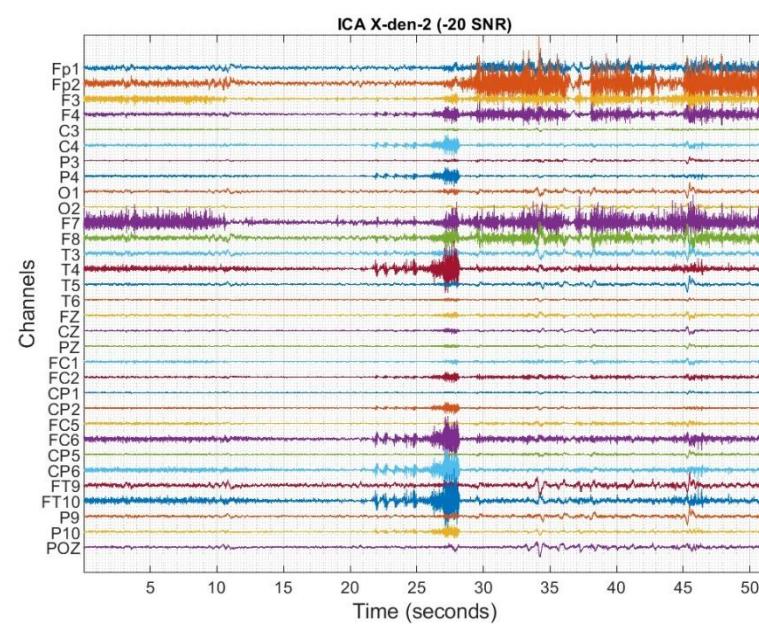
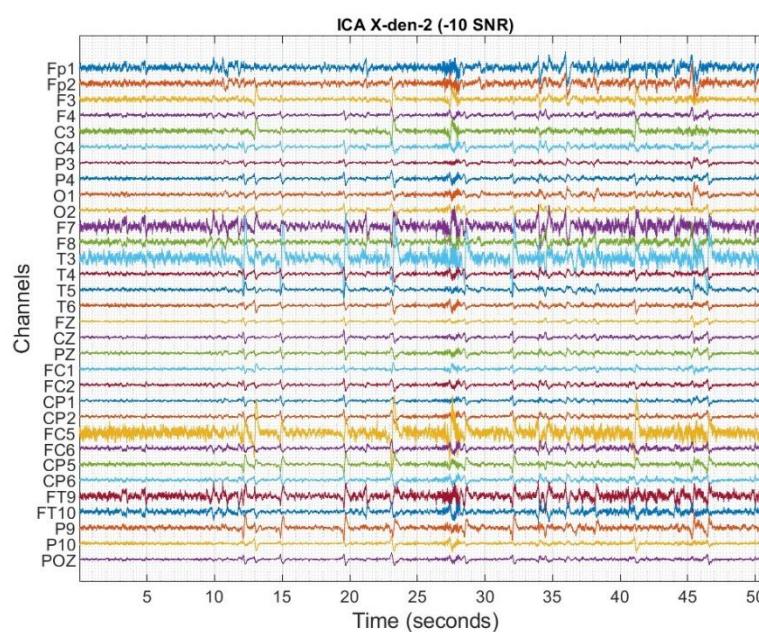
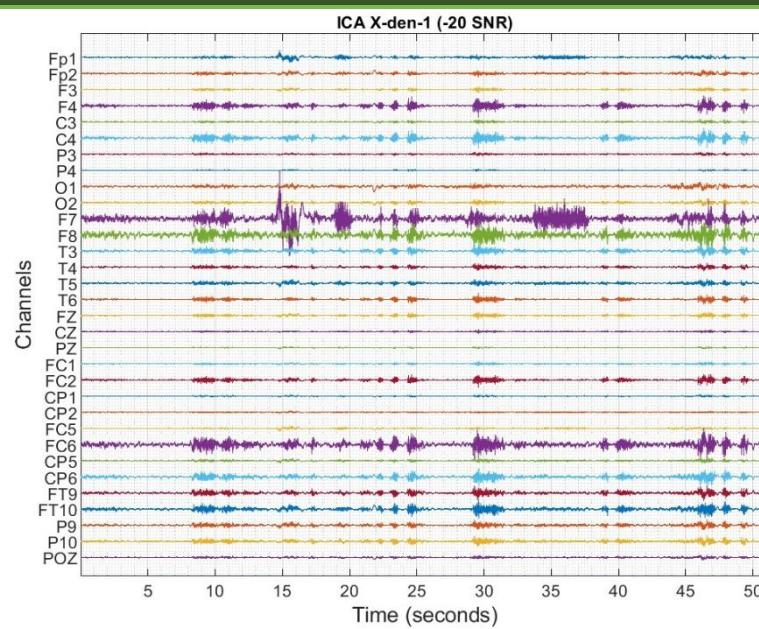
ج) با توجه به اسپایکی بودن منابع بدست آمده، اسپایکی ها را نگه می داریم و بقیه را حذف می کنیم. البته پیدا کردن منابع اسپایکی در PCA آسان نبود و ممکن است خطای زیادی داشته باشد. منابع نگه داشته شده به شرح زیر هستند.

```
SelSources_PCA_1 = [2 3 5 16 32];
SelSources_PCA_2 = [1 4 6 10 11 27 32];
SelSources_ICA_1 = [3 6 14 25];
SelSources_ICA_2 = [4 10 12 18 19];
```

د) حال با استفاده از منابع نگه داشته شده، مجددا سیگنال ۳۲ کاناله مشاهدات را برای هر دو روش و هر دو اس ان ارو هر دو نویز بدست می آوریم:



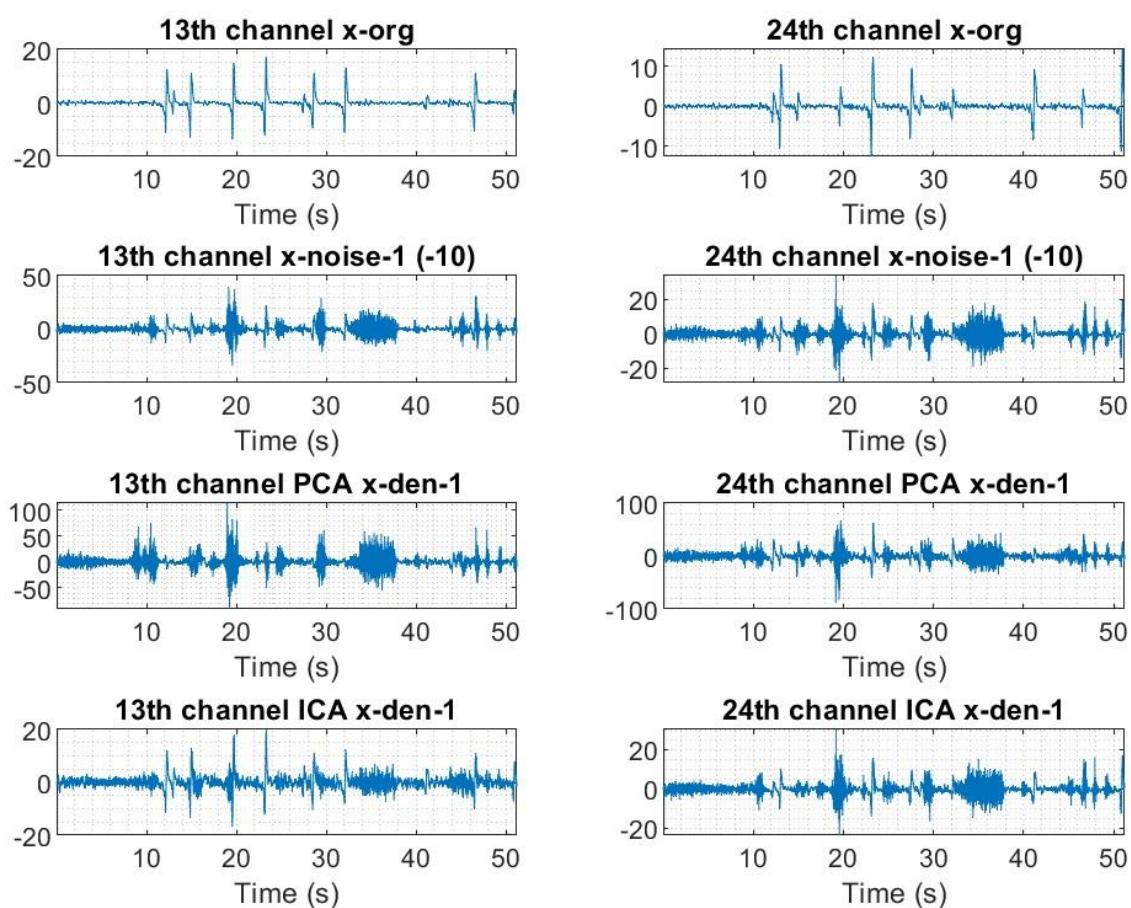




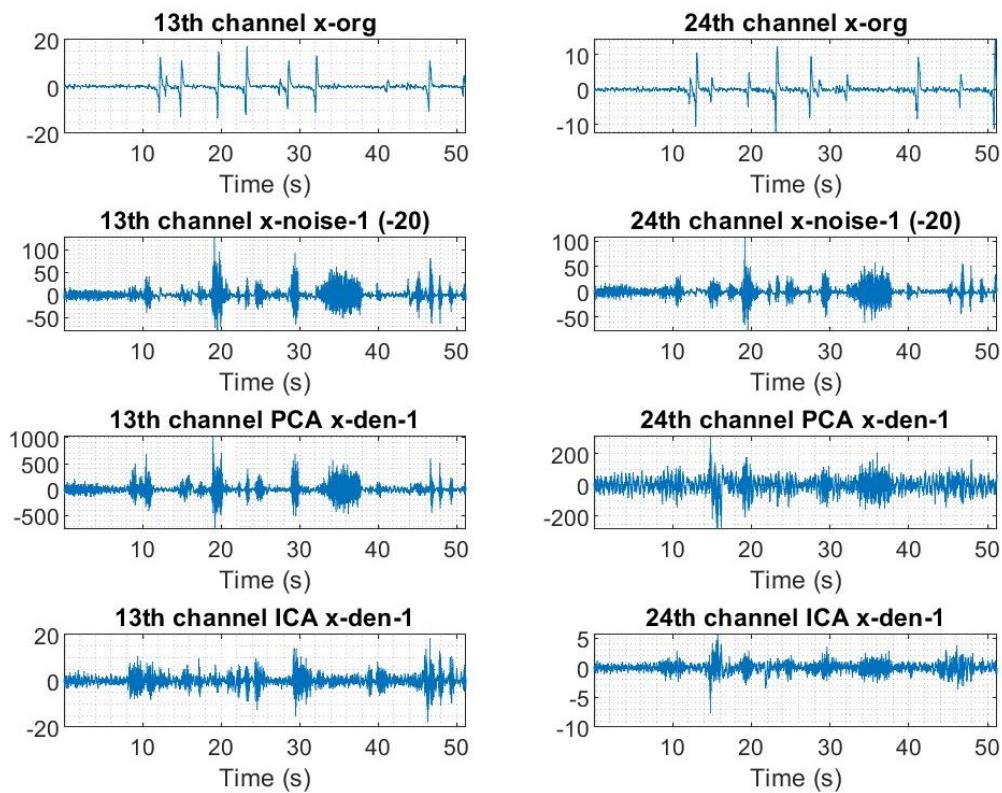
مشاهده می شود که در هر دو روش PCA و ICA با نگه داشتن سورس های اسپایکی و حذف دیگر سورس ها تاحدی توانسته ایم که حذف نویز کنیم و به سیگنال اصل نزدیک شویم. همانطور که انتظار داشتیم، روش ICA به نسبت عملکرد بهتری داشته است و به سیگنال اصلی نزدیک تر شده است. البته در هر دو روش اگر دقیق کنیم اسپایک ها در $\text{SNR} = -10$ بسیار معلوم تر از حالت دیگر هستند زیرا که توان نویز در آن ها کمتر است و حذف نویز بهتری در آن ها انجام شده است.

۵) در این بخش مشاهدات حذف نویز شده در هر دو روش را برای کانال های ۱۳ و ۲۴ رسم کرد و مقایسه می کنیم:

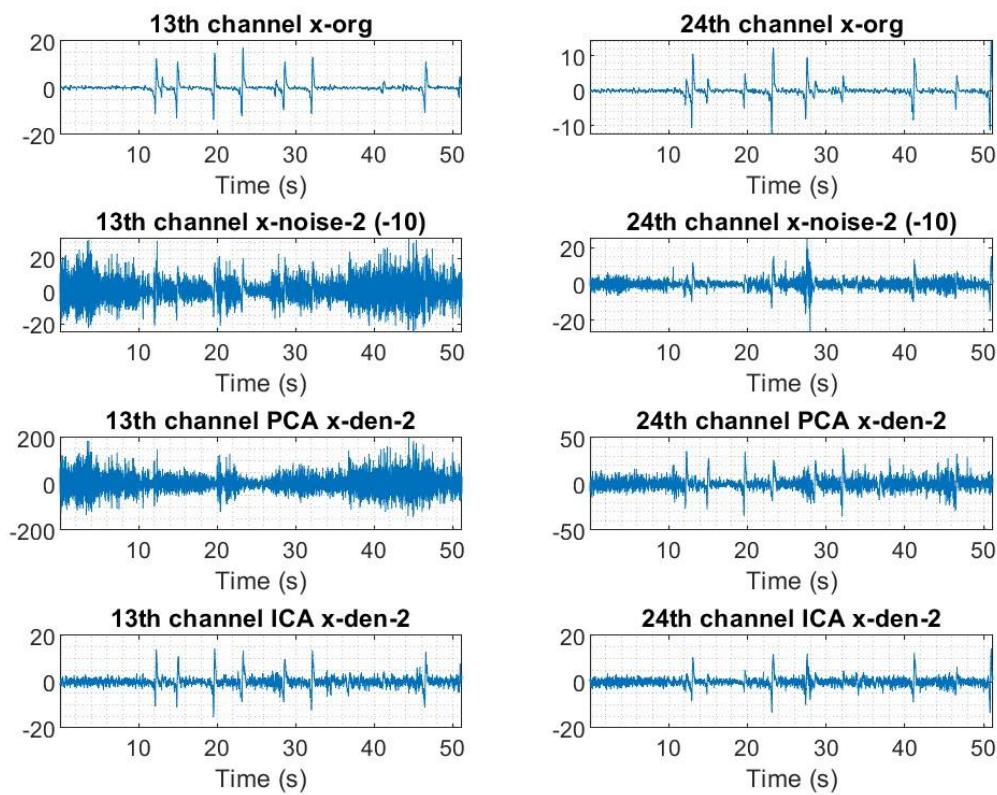
$$\text{نویز اول با } \text{SNR} = -10$$



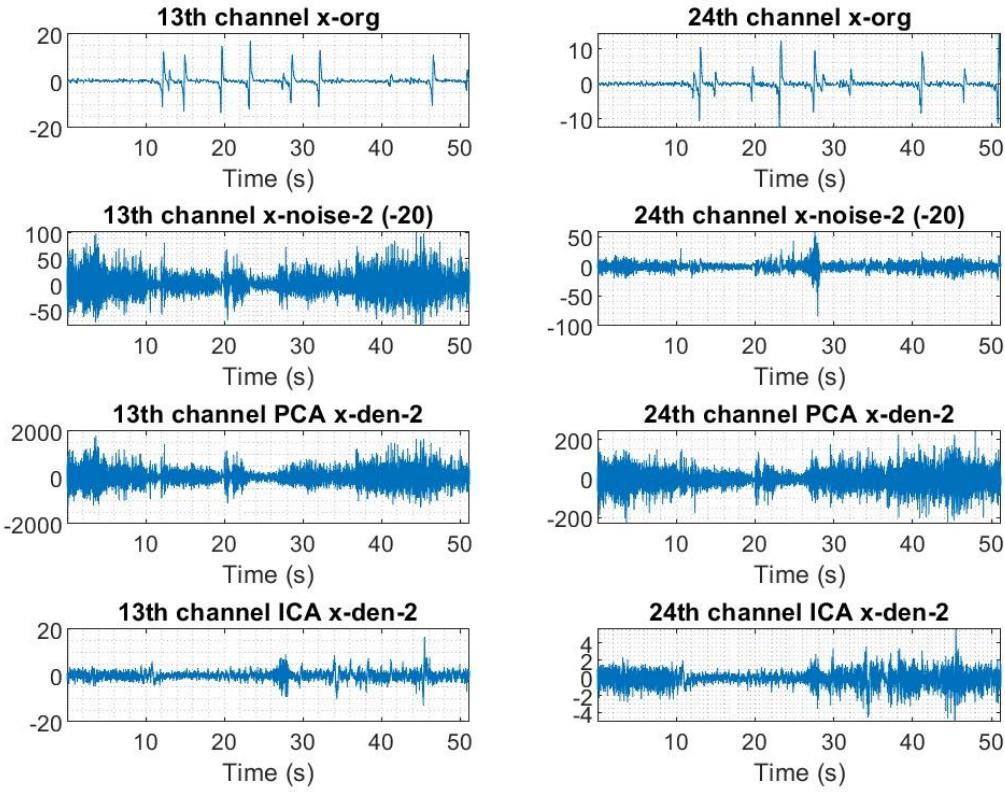
: SNR = -٢٠ با نویز اول



: SNR = -١٠ با نویز سوم



: SNR = -۲۰ با نویز سوم



نتایجی مشابه بحث های مطرح شده مشخص است. در اس ان ار های ۲۰- سیگنال های باز سازی شده شباهت کمتری به سیگنال اصلی دارند تا اس ان ار های ۱۰-. این مسئله درباره مزیت روش ICA نیز بر روی PCA صادق است. البته لزوما قرار نیست همیشه برقرار باشند، برای مثال می بینیم که در حالتی حذف نویز برای کانال ۱۳ بهتر عمل کرده است و حالی دیگر برای کانال ۲۴ عملکرد بهتری داشته است. به صورت کلی با روش های انجام شده به خصوص در نویز های با توان پایین تر توانستیم تا حدی حذف نویز را انجام دهیم و سیگنال بازسازی شده را به سیگنال ایده آل نزدیک تر کنیم.

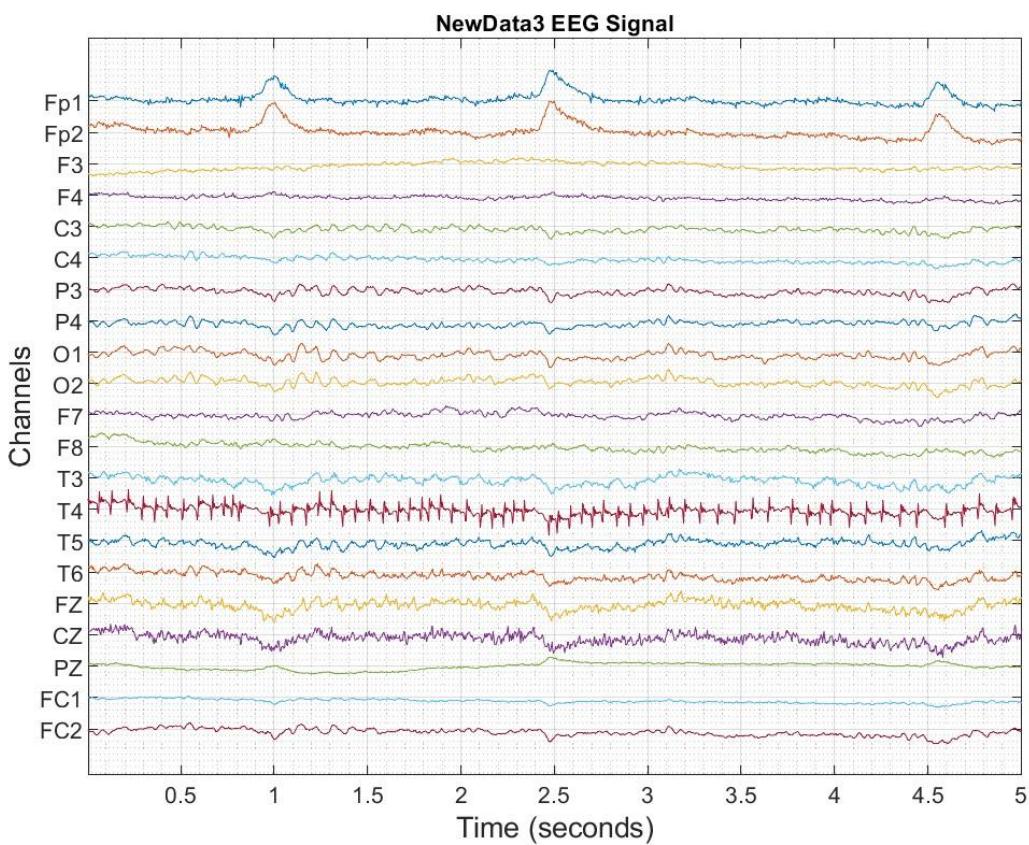
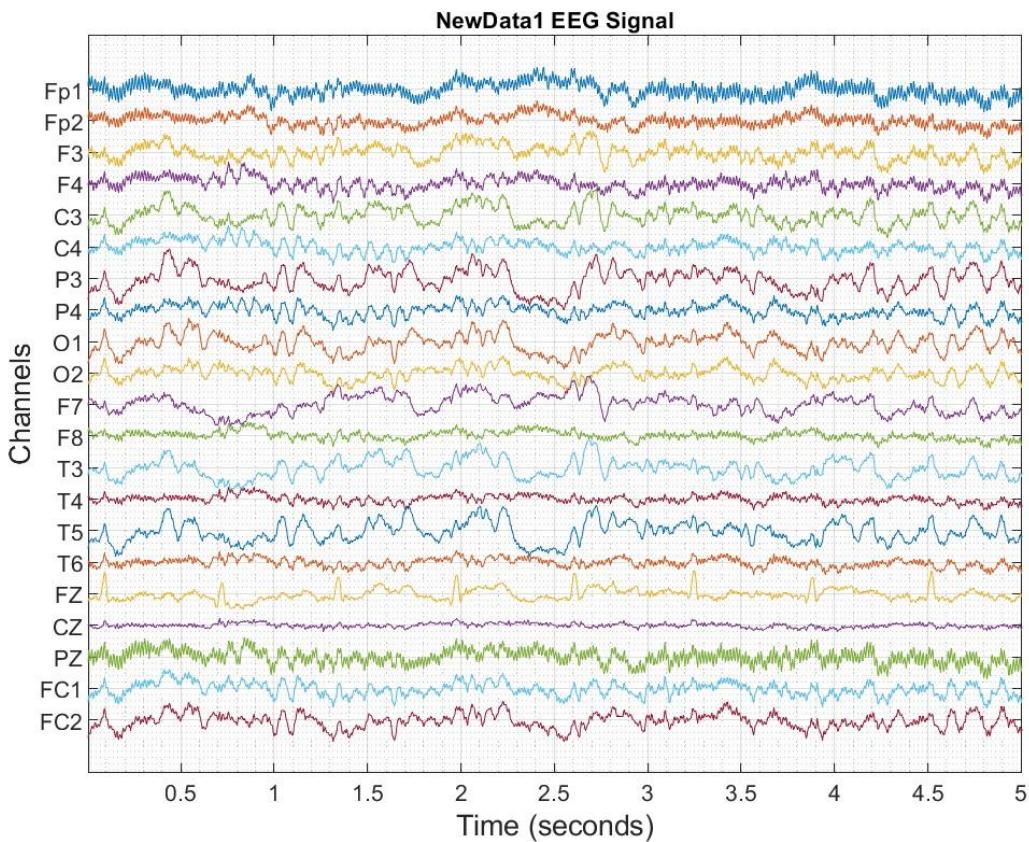
(و) مقدار خطای RRMSE را برای هر کدام حساب کرده و در جدول بیان می کنیم. توقع داریم که خطای در روش PCA به نسبت روش ICA بیشتر باشد. این توقع را نسبت به $-۲۰ = \text{SNR}$ نسبت به $-۱۰ = \text{SNR}$ می داریم.

	Noise ۱		Noise ۳	
	SNR = -۱۰	SNR = -۲۰	SNR = -۱۰	SNR = -۲۰
PCA	۶.۶۵۱۲	۶۱.۸۰۷۸	۱۰.۵۹۰۴	۱۰۰.۰۸۴۵
ICA	۱.۰۴۱۵	۴.۱۰۸۳	۰.۹۰۷۲	۳.۷۳۷۰

مقادیر بدست آمده بر سخنان بخش های قبل و این بخش مهر تایید می زند. روش PCA با کاهش SNR افت شدیدی پیدا می کند که این افت در ICA بسیار کمتر است.

سوال ۳:

الف) از فایل های اول و سوم استفاده می کنیم و آن ها را در زمان رسم می کنیم.



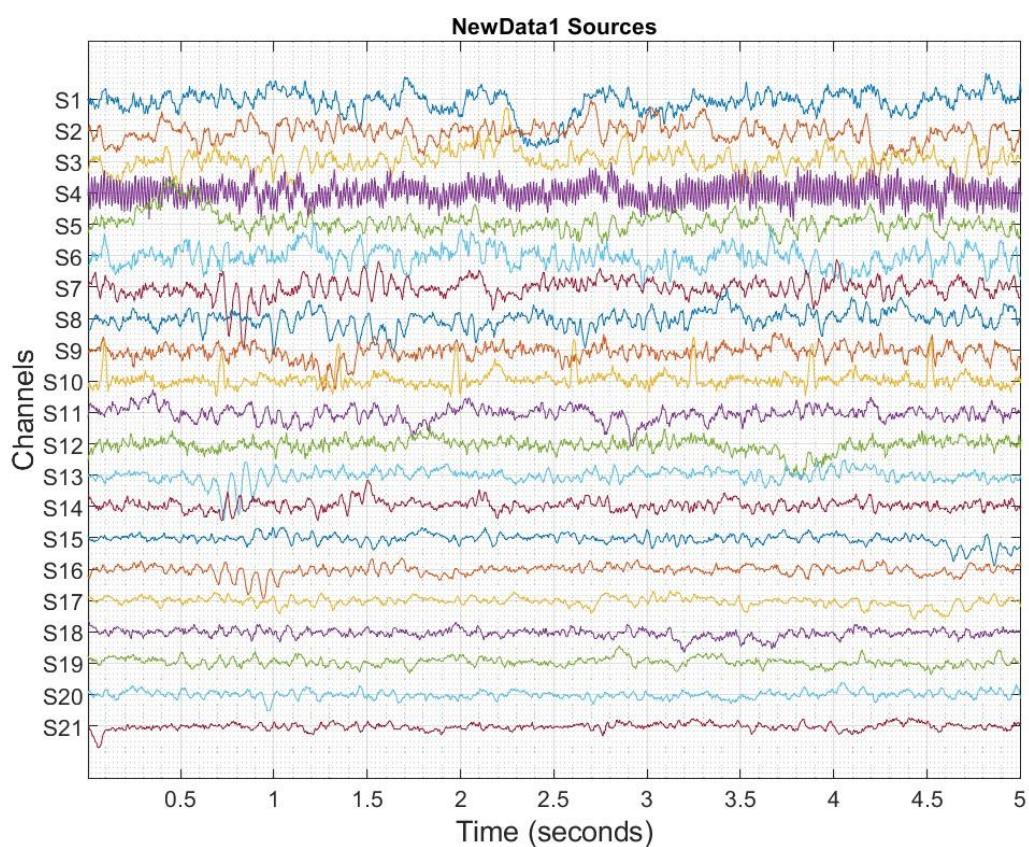
ب) در سیگنال اول می توان گفت که محتوای فرکانس های پایینتر است. و همچنین می توان گفت که نویز مشهود، نویز برق شهر است که بر روی سیگنال سوار است. اما در سیگنال دوم در دو تا از کanal ها پلک زدن مشاهده می شود. و در یکی از کanal ها نیز نویز ECG مشهود است. البته می توان گفت در یکی دیگر از کanal ها نویز ماهیچه ای نیز دیده می شود.

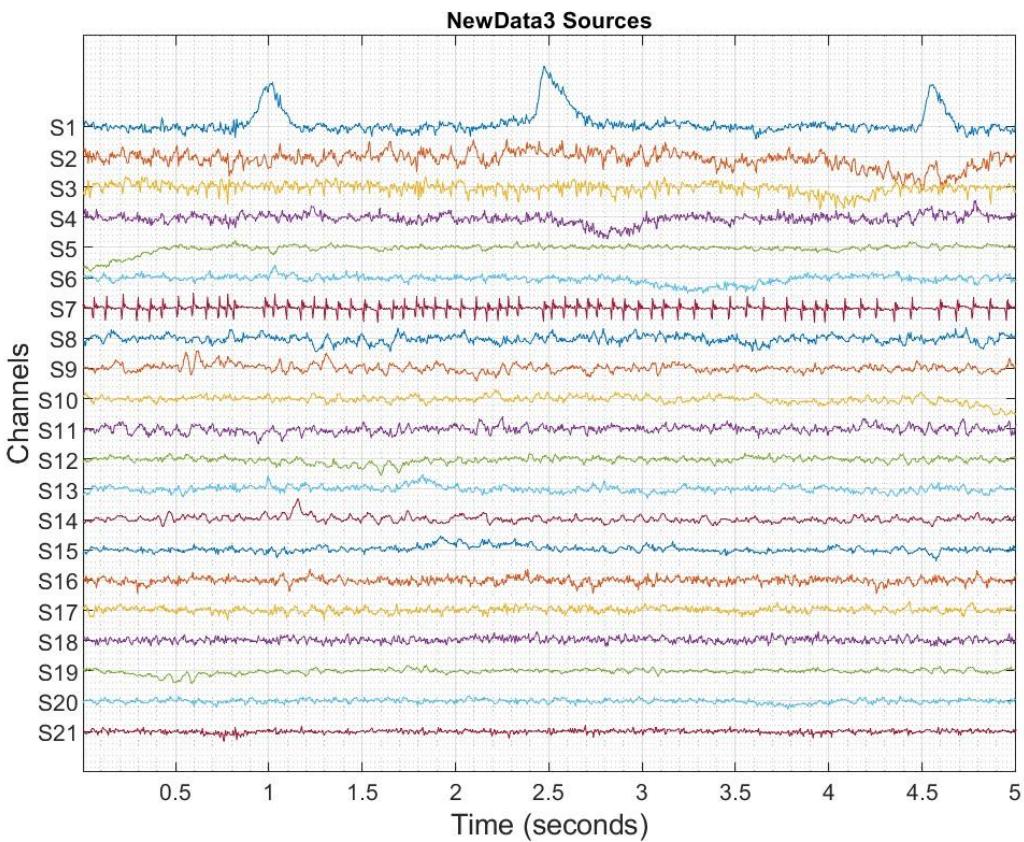
نویز های برق شهر، پلک زدن، ECG و همچنین نویز های عضلانی می توانند توسط ICA تا حد خوبی رفع شوند.

ج) با استفاده ازتابع COM2R الگوریتم ICA را بر روی دیتا ها پیاده می کنیم. W_1 و W_2 که ماتریس های ترکیب هستند را بدست می آوریم و سپس با استفاده از آن ها و دو سیگنال مشاهداتمان، مولفه ها را برای هر کدام بدست می آوریم : Sources_1 و Sources_2

د) می خواهیم که مولفه های بدست آمده را در حوزه های زمانی، فضایی و فرکانس بررسی کنیم.

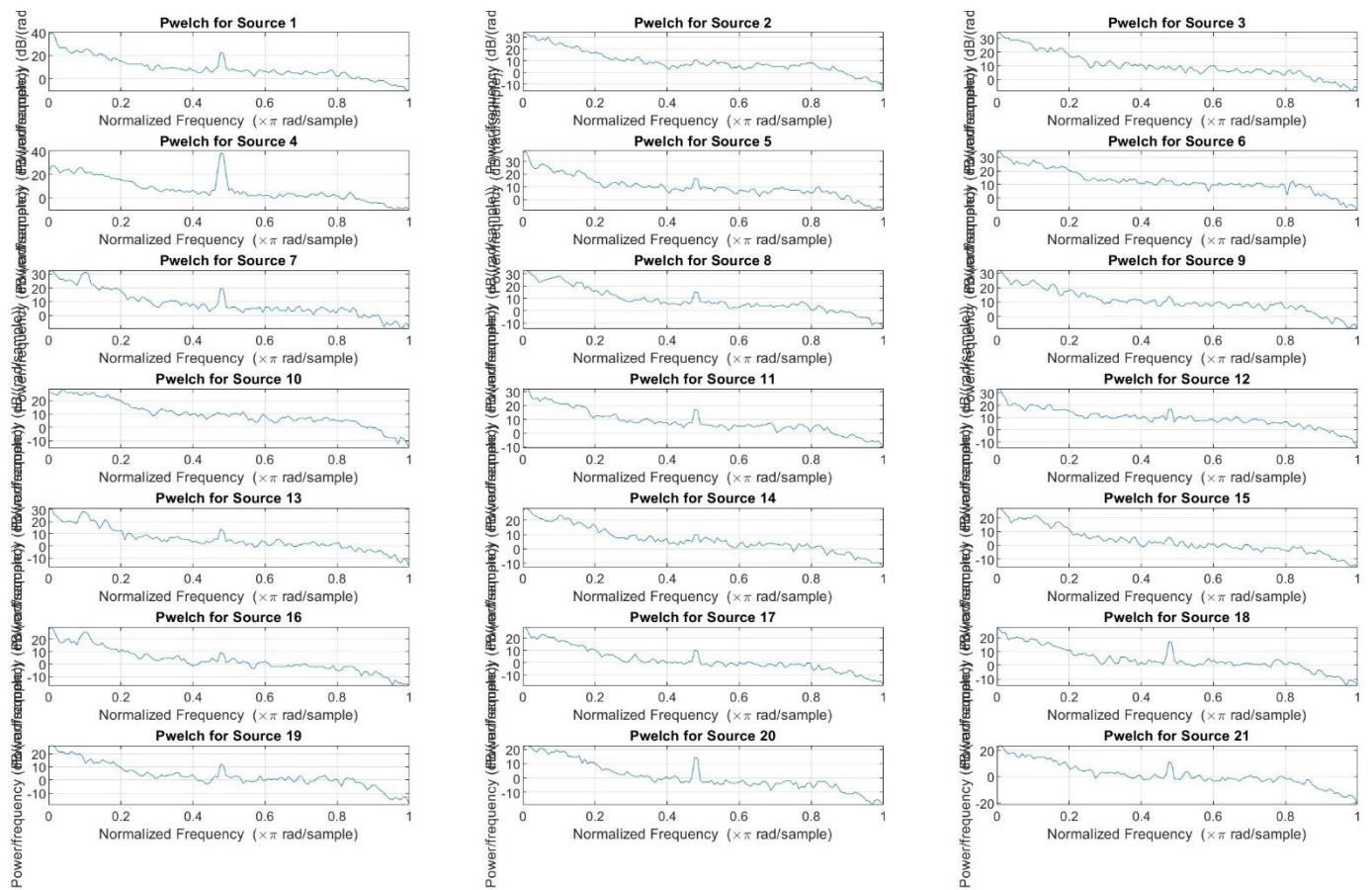
ابتدا سورس های بدست آمده را در زمان رسم می کنیم.



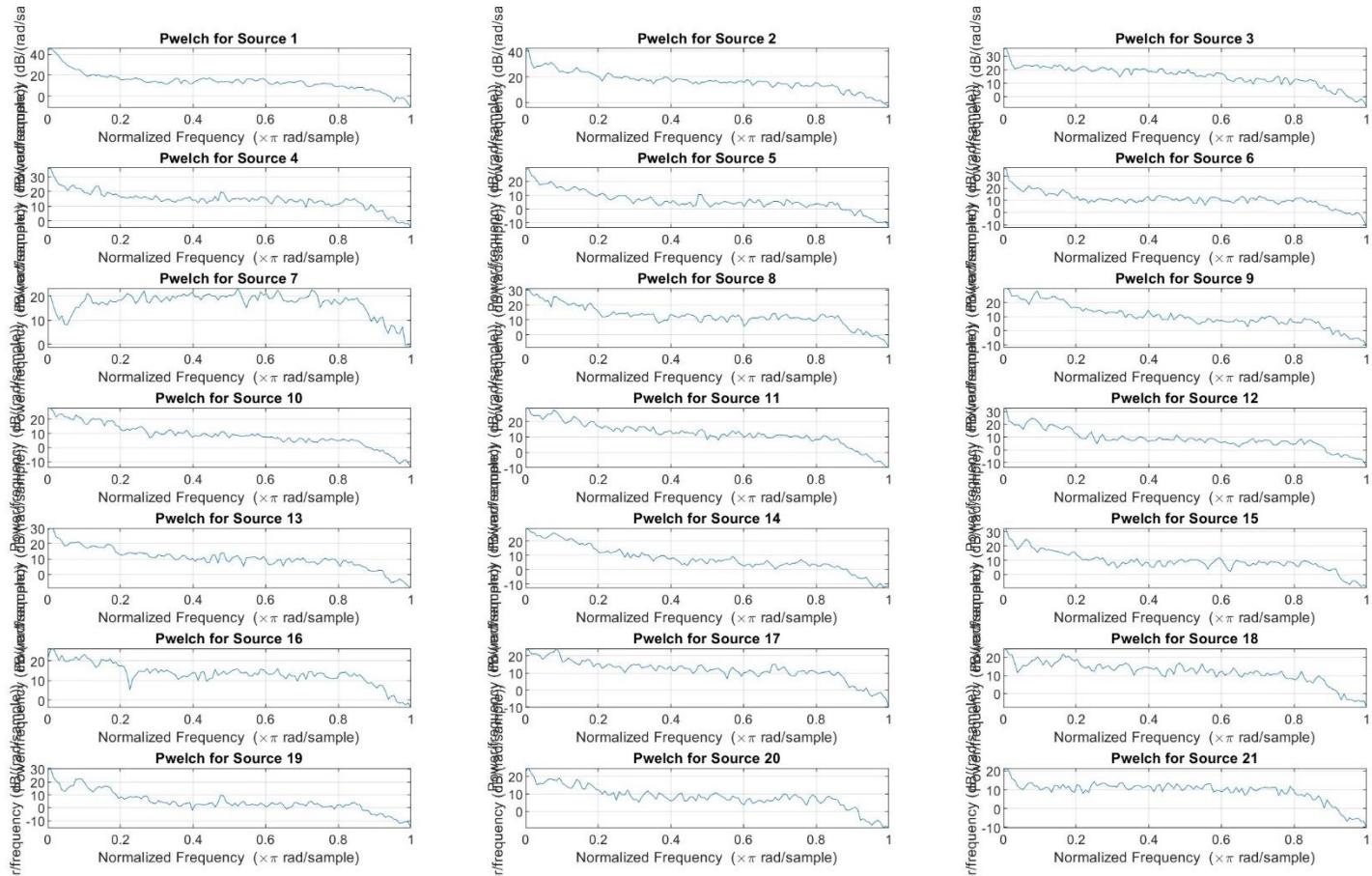


در دیتای اول نیاز است که مولفه ۴ را حذف کیم زیرا که احتمالاً نویز برق شهر را نشان می‌دهد. همچنین در دیتای سوم باید مولفه اول که حاصل پلک زدن است و مولفه ۷ که حاصل ECG است را حذف کیم.

حال به سراغ حوزه فرکانس می‌رویم، هر کدام از مولفه‌ها را به صورت جداگانه و سپس با هم با استفاده از تابع بررسی می‌کنیم. ابتدا منابع بدست آمده از دیتای اول را به صورت جداگانه می‌بینیم.

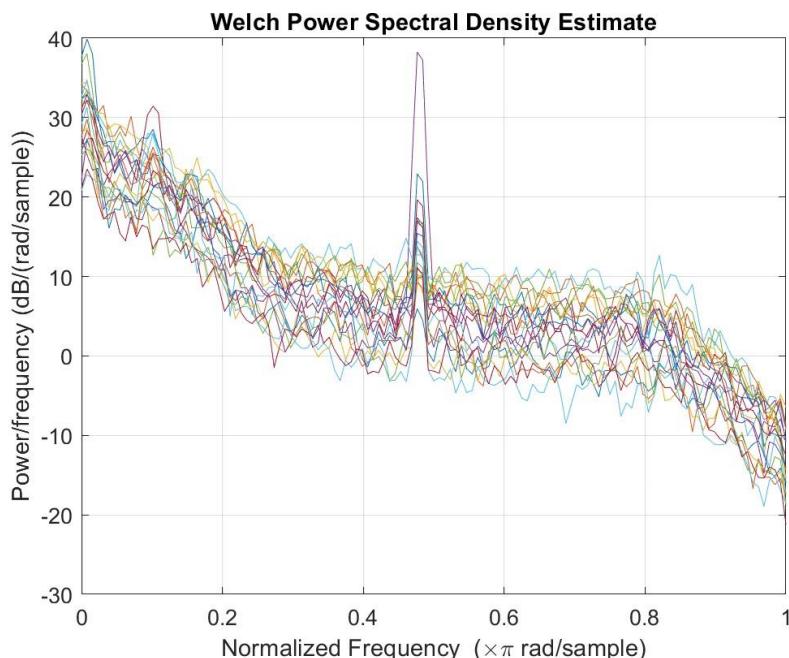


در این روش نیز مشهود است که مولفه ۴ حاوی دیتای زیادی از سمت نویز برق شهر است. البته مولفه های دیگری مانند شماره ۲۰ نیز این نویز را دارند اما در شماره ۴ این مسئله غالب تر است. برویم سراغ سیگنال سوم.

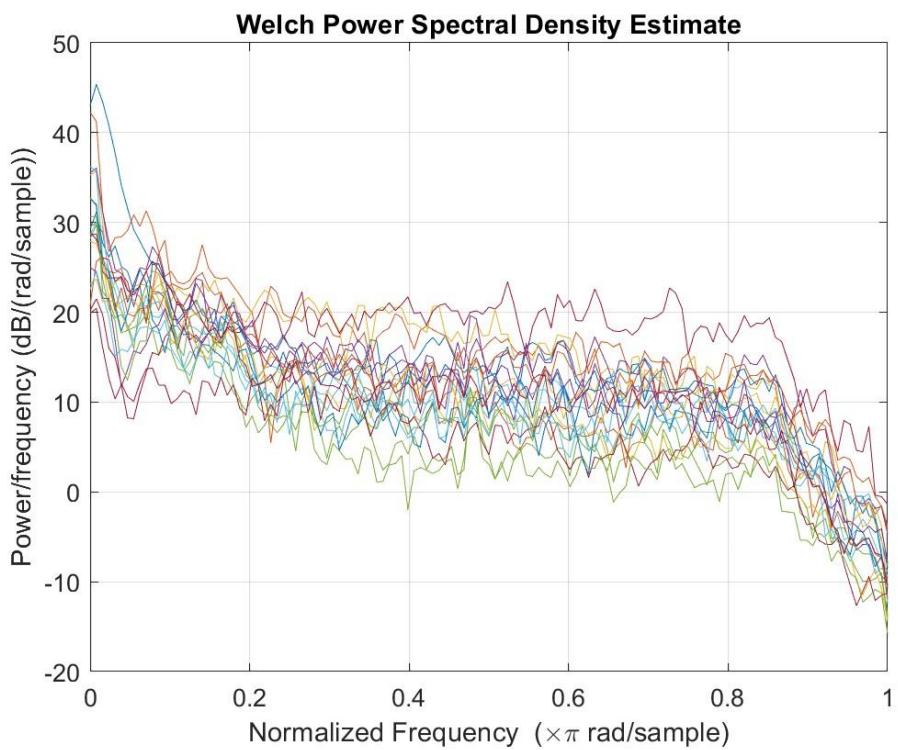


با استفاده از این دیتا نمی توان زیاد درباره پلک زدن صحبت کرد زیرا که فرکانسی بسیار کم است. اما همانطور که مشخص است، محتوای فرکانسی مولفه ۷ همانند حوزه زمان، با بقیه متفاوت است.

حال ها را روی هم بررسی می کنیم.

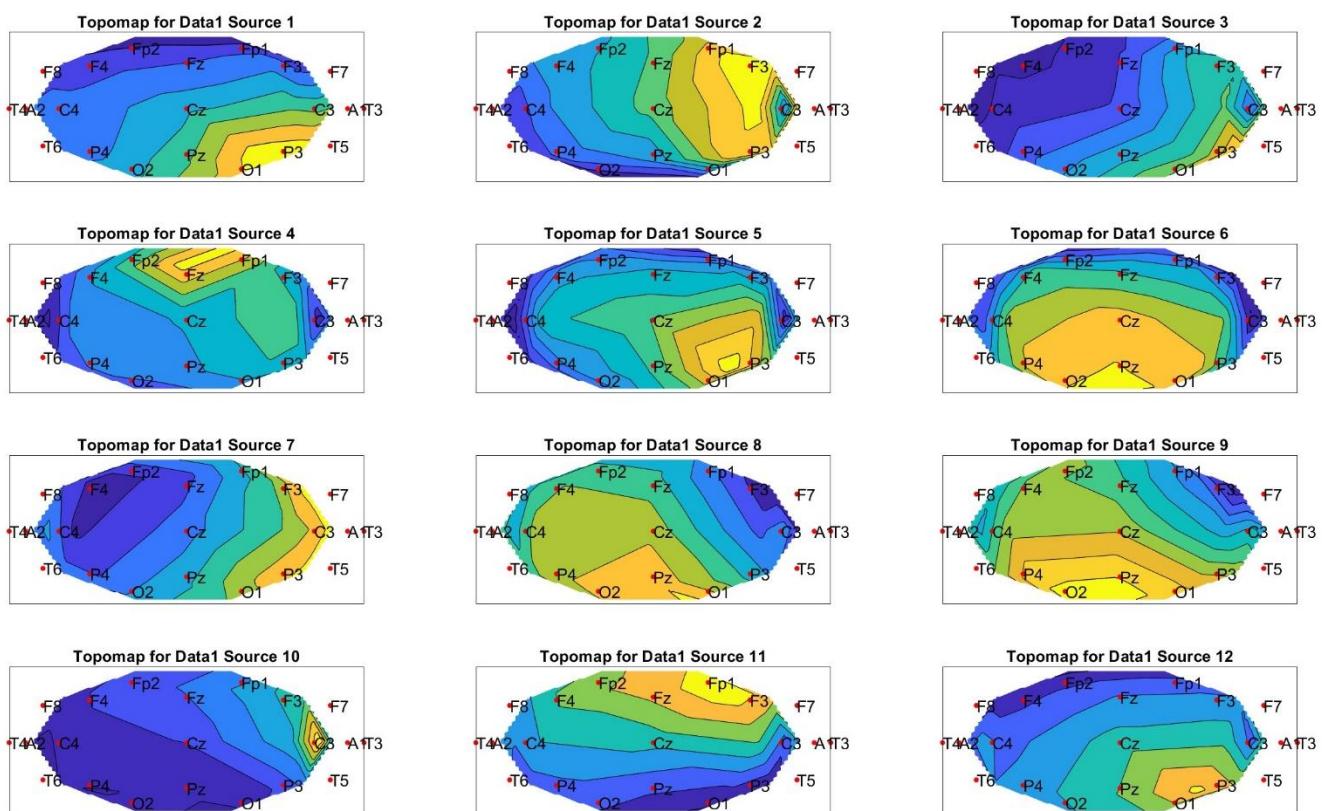


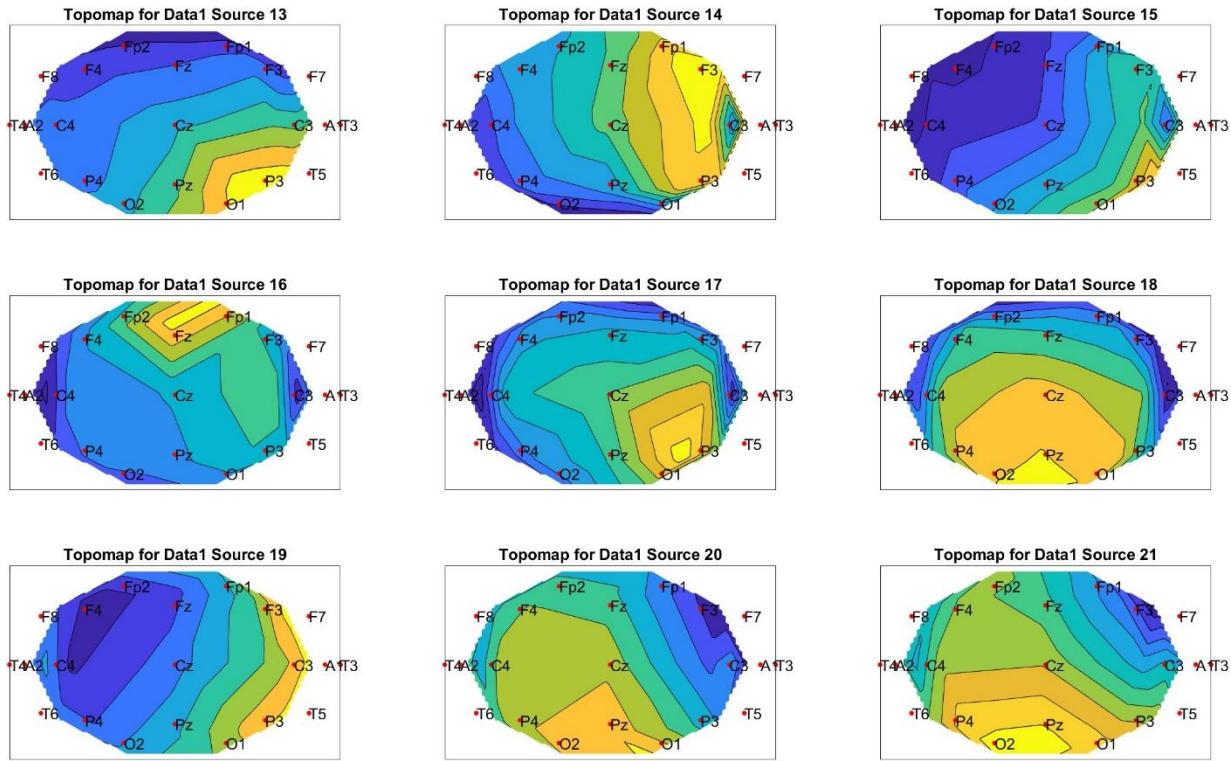
شاهد محتوای زیاد نویز برق در سیگنال ها هستیم که باید حذف شود. حال دیتای سوم:



خیلی قابلیت صحبت وجود ندارد اما می توان صرفا گفت که محتوای فرکانس های پایین در دیتای سوم از اولی بیشتر است.

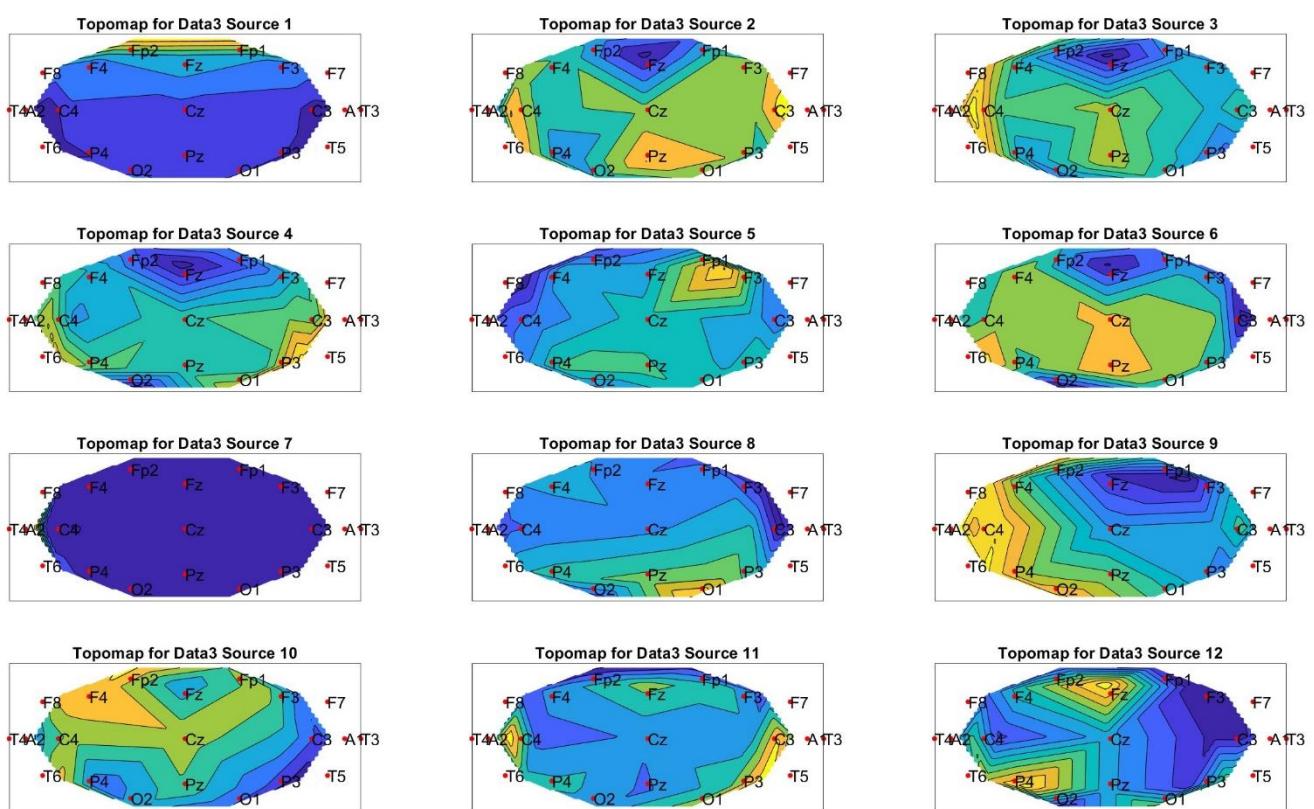
پس از بررسی در حوزه فرکانس، به سراغ بررسی در حوزه فضایی می رویم. ابتدا دیتای ۱:

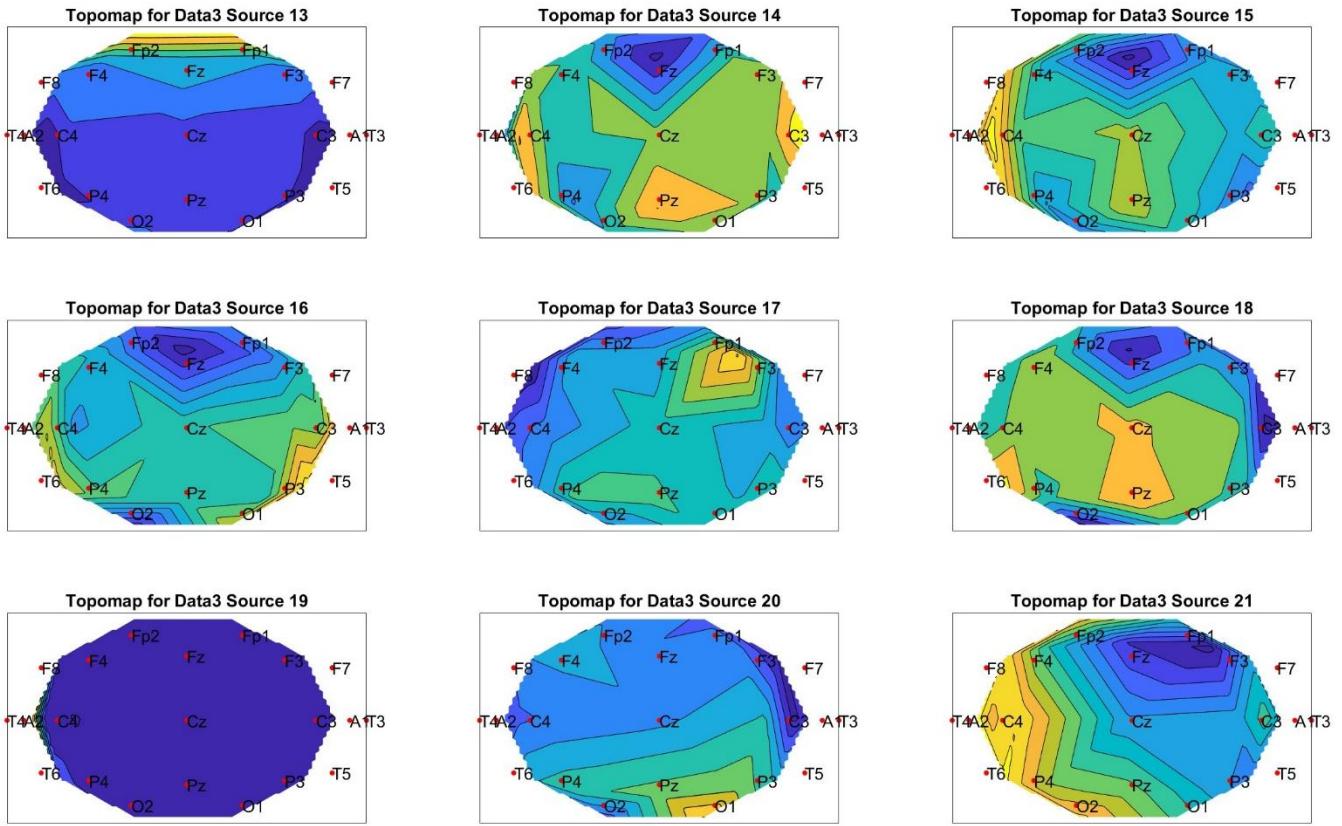




با توجه به توپوگرافی مشاهده شده احتمالاً سورس شماره ۱۰ حاصل نویز ECG یا EMG است. حال که با دقت به دیتای حوزه زمان نگاه میکنیم مشخص می شود که بله ECG است و باید حذف شود.

حال به سراغ دیتای سوم می رویم:





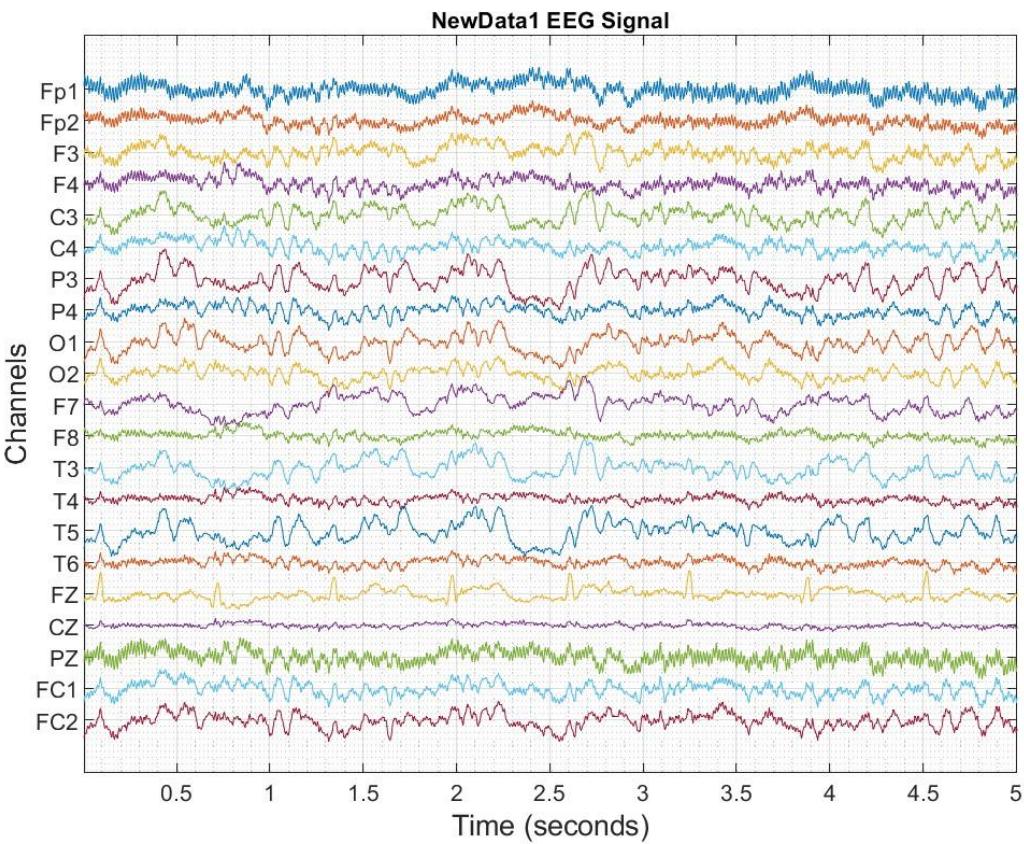
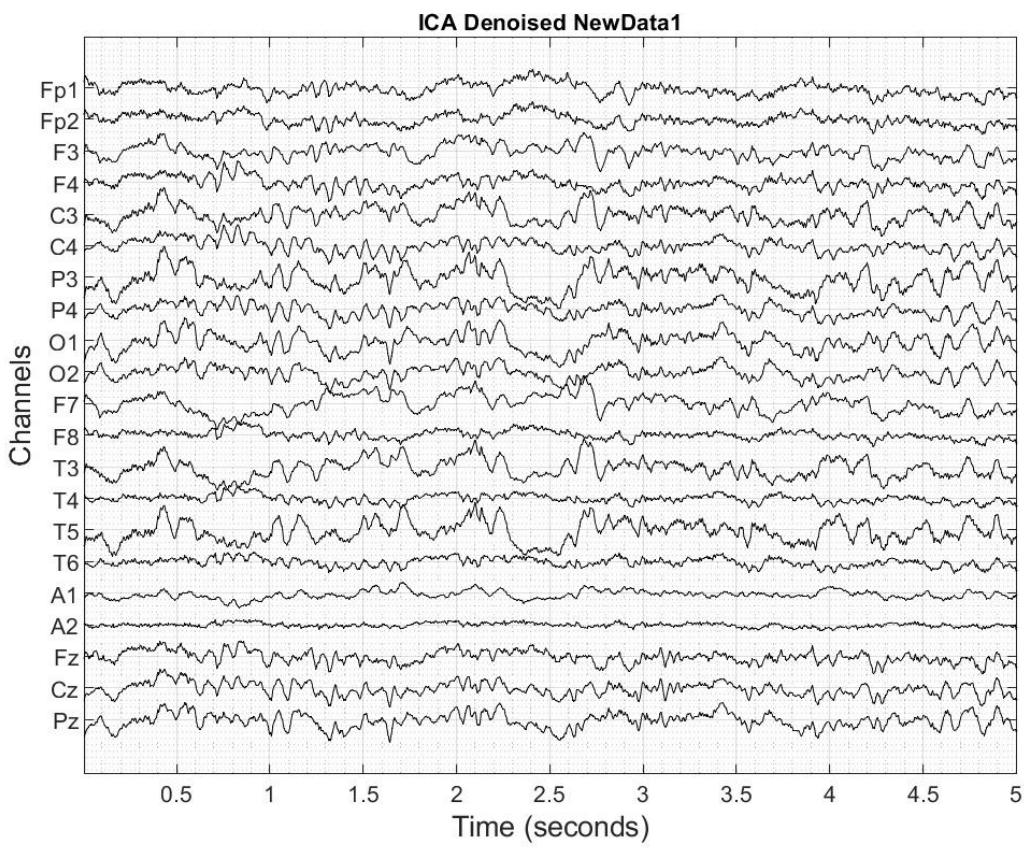
برای این دیبا توقع می رود که اطلاعات بهتری بدست آید. همانطور که توقع داشتیم مطمئن می شویم که سورس اول ناشی از نویز پلک زدن است و می توانیم آن را حذف کنیم. سورس شماره ۷ نیز گویا یا ECG و یا نوعی پرش پلک یا مشابه آن است که توپوگرافی خیلی دقیقی دارد و باید حذف شود زیرا که سیگنال مغزی نیست. می توان درباره سورس ۱۳ نیز مشکوک بود که امکان دارد برای حرکت چشم باشد. اما ترجیح بر نگه داشتن است. سورس شماره ۱۹ نیز با توجه به توپوگرافی بسیار مشکوک است و می توانیم آن را حذف کنیم.

۵) از دیتای اول سورس های شماره ۴ و ۱۰ را حذف می کنیم و از دیتای شماره سه نیز سورس های ۱ و ۷ و ۱۹ را حذف می کنیم. سپس سیگنال را بازسازی می کنیم.

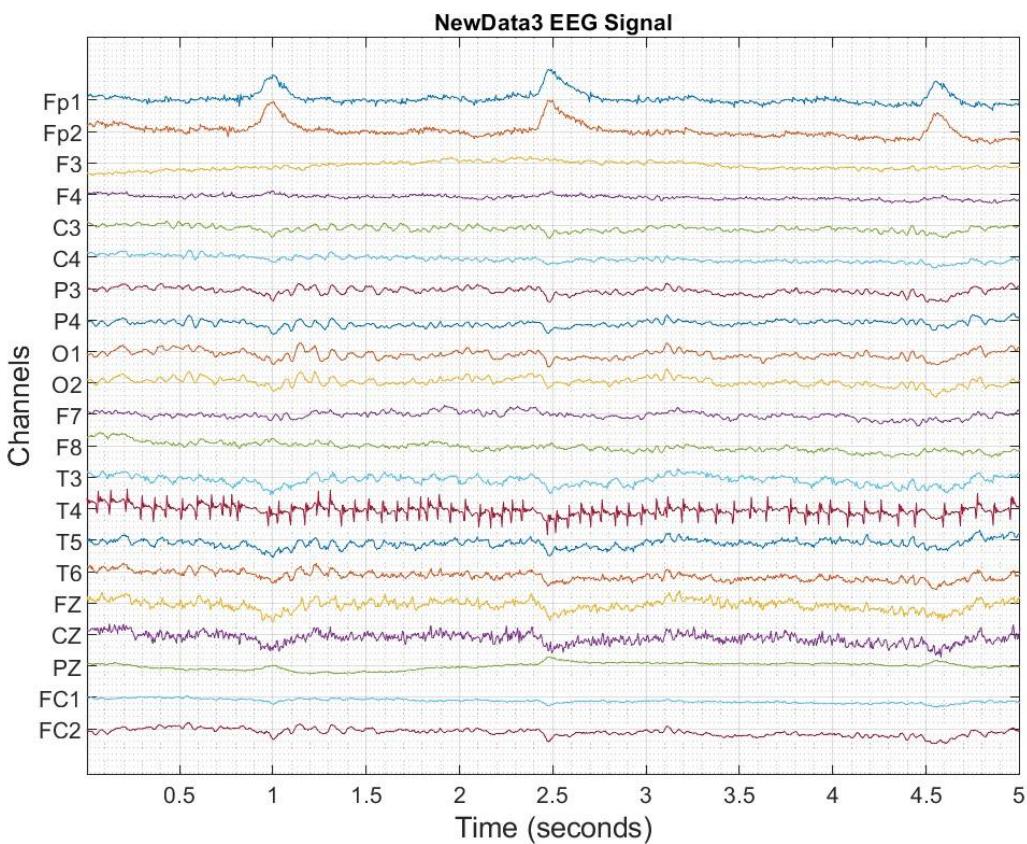
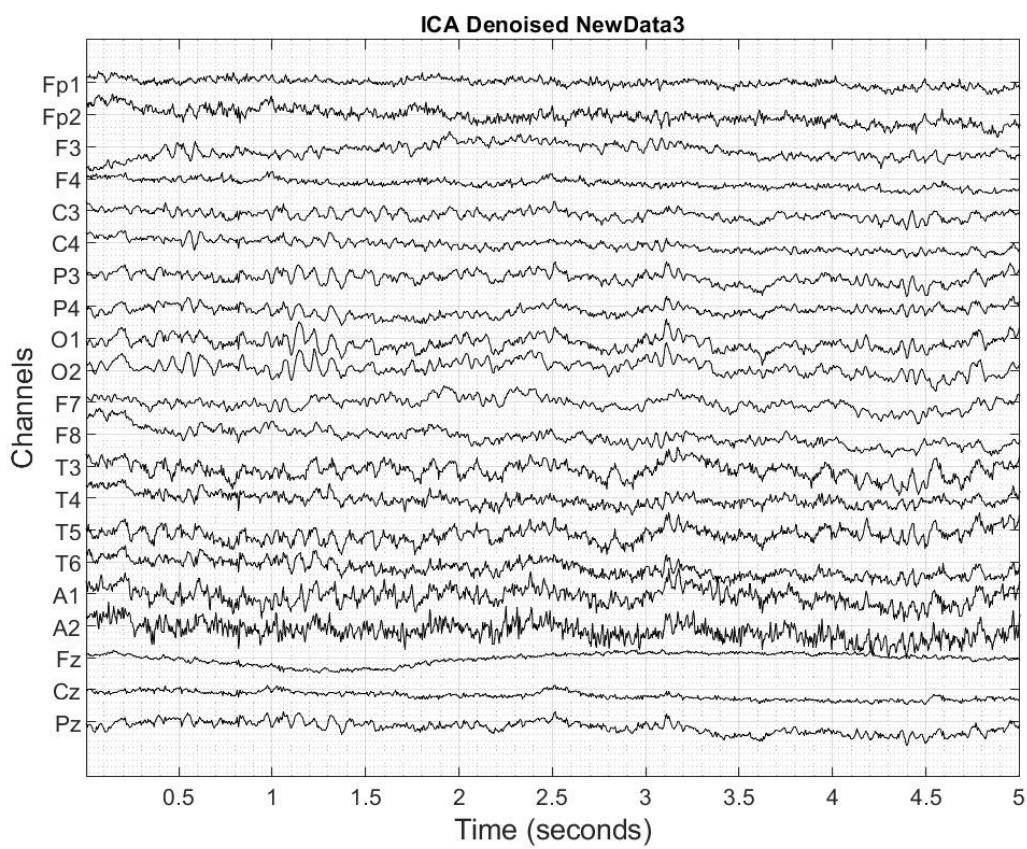
با توجه به حذف نویز های اعمال شده، به نظر می آید که نیازی به تکرار مرحله حذف نویز نیست.

در دیتای اول برق شهر و ECG به خوبی حذف شده است و سیگنال نرم تر شده است.

در دیتای دوم نیز سیگنال پلک زدن و نویز ECG و همچنین نویز اندر EMG حذف شده و دیگر شاهد سیگنال های غالب با دامنه بزرگ فرکانس پایین ناشی از پلک زدن نیستیم.



به صورت کلی نویز ناشی از برق شهر حذف شده است و همچنین در FZ دیگر نویز ECG دیده نمی شود.



نویز پلک زدن از دو کانال اول حذف شده و همچنین نویز موجود در T4 نیز رفع شده است.