Supplemental materials to "Global Testing Against Sparse Alternatives in Time-Frequency Analysis"

0.1. *Proof of Lemmas 5.2.*

PROOF. Since the density of z is $\frac{1}{\pi}e^{-|z|^2}$, we have

$$\bar{\Psi}(t) = \int_{|z|>t} \frac{1}{\pi} e^{-|z|^2} dA,$$

where dA is the Lebesgue measure on the z-plane. In other words, if $z = x + iy$, then $dA = dxdy$. By applying the polar coordinates $z = re^{i\theta}$, we have

$$\bar{\Psi}(t) = \int_0^{2\pi} \int_t^{\infty} \frac{1}{\pi} e^{-r^2} r dr d\theta = \int_t^{\infty} 2re^{-r^2} dr = e^{-t^2}.$$

To prove (5.1), define $u \in \mathbb{C}$ satisfying $|u| = 1$ and $\bar{u}\mu = |\mu|$. This unit complex scalar always exists since we can let $u = \frac{\mu}{|\mu|}$ when $\mu \neq 0$, and any unit scalar when $\mu = 0$. Notice that

$$\Re(\bar{u}z) > t - |\mu| \implies \Re(\bar{u}z) + |\mu| > t \implies \Re(\bar{u}(z+\mu)) > t \implies |z+\mu| > t,$$

and hence

$$\mathbb{P}(|z+\mu| > t) \geq \mathbb{P}(\Re(\bar{u}z) > t - |\mu|).$$

Since

$$z \sim \mathcal{CN}(0,1,0) \implies \bar{u}z \sim \mathcal{CN}(0,1,0) \implies \Re(\bar{u}z) \sim \mathcal{N}\left(0, \frac{1}{2}\right),$$

by the tail probability of standard real-valued normal variable we have

$$\mathbb{P}(\Re(\bar{u}z) > t - |\mu|) \geq \frac{C_0}{1 + (t-|\mu|)_+} e^{-(t-|\mu|)_+^2}.$$

Moreover,

$$\mathbb{P}(|\mu + z| > t) \leq \mathbb{P}(|z| > t - |\mu|) \leq e^{-(t-|\mu|)_+^2}.$$

□

0.2. *Proof of Lemma 5.3.*

PROOF. Simple calculation yields

(0.13)
$$\mathrm{Cov}(1_{\{|w_1-a_1|>t\}}, 1_{\{|w_2-a_2|>t\}}) = \mathbb{P}(|w_1 - a_1| > t, |w_2 - a_2| > t) - \mathbb{P}(|w_1 - a_1| > t)\,\mathbb{P}(|w_2 - a_2| > t).$$

This implies

$$\begin{aligned}\mathrm{Cov}(1_{\{|w_1-a_1|>t\}},1_{\{|w_2-a_2|>t\}}) &\le \mathbb{P}(|w_1-a_1|>t,|w_2-a_2|>t)\\ &\le \min\left(\mathbb{P}(|w_1-a_1|>t),\mathbb{P}(|w_2-a_2|>t)\right)\\ &\le \min\left(\mathbb{P}(|w_1|>(t-|a_1|)_+),\mathbb{P}(|w_2|>(t-|a_2|)_+)\right)\\ &= \min\left(e^{-(t-|a_1|)_+^2},e^{-(t-|a_2|)_+^2}\right).\end{aligned}$$

When $|\xi|\le\frac{1}{2}$, let $\mathbf{\Gamma}_h=\begin{bmatrix}1 & h\xi\\ h\bar{\xi} & 1\end{bmatrix}$, $0\le h\le 1$. Define a random vector $\begin{bmatrix}Z_1\\ Z_2\end{bmatrix}\sim\mathcal{CN}(0,\mathbf{\Gamma}_h,\mathbf{0})$ and a function of h

$$F(h)=\mathbb{P}(|Z_1-a_1|>t,|Z_2-a_2|>t).$$

Here $t>0$ is a fixed parameter. By Newton-Lebnitz theorem, we have

$$\mathrm{Cov}(1_{\{|w_1-a_1|>t\}},1_{\{|w_2-a_2|>t\}})=F(1)-F(0)=\int_0^1 F'(h)dh.$$

It suffices to give an upper bound to $F'(h)$ for all $0<h<1$. By the density function of $\begin{bmatrix}Z_1\\ Z_2\end{bmatrix}$, we have the explicit formula

$$F(h)=\int\limits_{|z_1-a_1|>t}\int\limits_{|z_2-a_2|>t}\frac{1}{\pi^2\det(\mathbf{\Gamma}_h)}\exp\left(-\begin{bmatrix}z_1\\ z_2\end{bmatrix}^*\mathbf{\Gamma}_h^{-1}\begin{bmatrix}z_1\\ z_2\end{bmatrix}\right)dA_2dA_1,$$

Here dA_1 and dA_2 are the Lebesgue measures on the z_1-plane and z_2-plane, respectively. In other words, if we write $z_1=x_1+iy_1$ and $z_2=x_2+iy_2$, then $dA_1=dx_1dy_1$ and $dA_2=dx_2dy_2$. Simple calculation in linear algebra gives $\det(\mathbf{\Gamma}_h)=1-h^2|\xi|^2$ and $\mathbf{\Gamma}_h^{-1}=\frac{1}{1-h^2|\xi|^2}\begin{bmatrix}1 & -h\xi\\ -h\bar{\xi} & 1\end{bmatrix}$, which implies

$$F(h)=\iint\limits_{\substack{|z_1-a_1|>t\\ |z_2-a_2|>t}}\frac{1}{\pi^2(1-h^2|\xi|^2)}\exp\left(-\frac{|z_1|^2+|z_2|^2-2h\Re(\xi\bar{z}_1z_2)}{1-h^2|\xi|^2}\right)dA_2dA_1.$$

By changing the order of derivative and integrals, we have

$$
\begin{aligned}
&F'(h)\\
&= \iint\limits_{\substack{|z_1-a_1|>t\\|z_2-a_2|>t}} \exp\left(-\frac{|z_1|^2+|z_2|^2-2h\Re(\xi\bar{z}_1z_2)}{1-h^2|\xi|^2}\right)\\
&\left[\frac{2h|\xi|^2}{\pi^2(1-h^2|\xi|^2)^2}+\frac{2\Re(\xi\bar{z}_1z_2)+(2h^2\Re(\xi\bar{z}_1z_2)-2(|z_1|^2+|z_2|^2)h)|\xi|^2}{\pi^2(1-h^2|\xi|^2)^3}\right]dA_2dA_1\\
&\le C|\xi| \iint\limits_{\substack{|z_1-a_1|>t\\|z_2-a_2|>t}} (1+|z_1||z_2|)\exp\left(-\frac{|z_1|^2+|z_2|^2-2h\Re(\xi\bar{z}_1z_2)}{1-h^2|\xi|^2}\right)dA_2dA_1,
\end{aligned}
$$

where C is a numerical constant. The last inequality is due to $0\le h\le 1$ and $|\xi|\le\frac{1}{2}$. Notice that

$$
\begin{aligned}
\frac{|z_1|^2+|z_2|^2-2h\Re(\xi\bar{z}_1z_2)}{1-h^2|\xi|^2} &\ge \frac{|z_1|^2+|z_2|^2-2h|\xi||z_1||z_2|}{(1-h|\xi|)(1+h|\xi|)}\\
&\ge \frac{|z_1|^2+|z_2|^2}{1+h|\xi|}.
\end{aligned}
$$

This is due to the fact that $\frac{|z_1|^2+|z_2|^2-2\rho|z_1||z_2|}{(1-\rho)}$ obtains the minimum at $\rho=0$. Therefore

$$
\begin{aligned}
F'(h) &\le C|\xi| \iint\limits_{\substack{|z_1-a_1|>t\\|z_2-a_2|>t}} (1+|z_1||z_2|)\exp\left(-\frac{|z_1|^2+|z_2|^2}{1+h|\xi|}\right)dA_2dA_1\\
&\le C|\xi| \iint\limits_{\substack{|z_1|>(t-|a_1|)_+\\|z_2|>(t-|a_2|)_+}} (1+|z_1||z_2|)\exp\left(-\frac{|z_1|^2+|z_2|^2}{1+|\xi|}\right)dA_2dA_1.
\end{aligned}
$$

By using the polar coordinates: $z_1=r_1e^{i\theta_1}$ and $z_2=r_2e^{i\theta_2}$, we have $dA_1=r_1dr_1d\theta_1$ and $dA_2=r_2dr_2d\theta_2$. Then

$$
F'(h)\le C|\xi|4\pi^2\int_{(t-|a_1|)_+}^{\infty}\int_{(t-|a_2|)_+}^{\infty}(1+r_1r_2)\exp\left(-\frac{r_1^2+r_2^2}{1+|\xi|}\right)r_1r_2dr_1dr_2.
$$

For any fixed $u>0$, simple integration by parts yields

$$
\int_u^\infty \exp\left(-\frac{r^2}{1+|\xi|}\right)rdr=\frac{1+|\xi|}{2}\exp\left(-\frac{u^2}{1+|\xi|}\right),
$$

and

$$\begin{aligned}
&\int_u^\infty \exp\left(-\frac{r^2}{1+|\xi|}\right) r^2 dr \\
&= \frac{(1+|\xi|)u}{2}\exp\left(-\frac{u^2}{1+|\xi|}\right) + \frac{1+|\xi|}{2}\int_u^\infty \exp\left(-\frac{r^2}{1+|\xi|}\right) dr \\
&\le C(1+u)\exp\left(-\frac{u^2}{1+|\xi|}\right),
\end{aligned}$$

where the last inequality is due to the real Gaussian bound. These equalities/inequalities give

$$F'(h) \le C_0|\xi| \exp\left(-\frac{(t-|a_1|)_+^2 + (t-|a_2|)_+^2}{(1+|\xi|)}\right)(1+(t-|a_1|)_+)(1+(t-|a_2|)_+).$$

By integrating it over $[0,1]$, our proposition is proven. □

0.3. *Pproof of Theorem 2.2.*

PROOF. Without loss of generality, assume $\frac{p}{N} = p^\gamma$ is an integer. We now consider a class of special alternatives. Let

$$\tilde\beta_1 = \ldots = \tilde\beta_s = \sqrt{\frac{rp\log p}{N}}$$

be real and positive. As to $\boldsymbol\tau$, we first define a set of index vectors:

$$\begin{aligned}
T_N = \Big\{ &\tilde{\boldsymbol\tau} = (\tilde\tau_1, \ldots, \tilde\tau_s) : 1 \le \tilde\tau_1 < \ldots < \tilde\tau_s \le N, \\
&\tilde\tau_{l+1} - \tilde\tau_l \ge \log^2 N \text{ for } l = 1, \ldots, s-1, \tilde\tau_1 + N - \tilde\tau_s \ge \log^2 N \Big\}.
\end{aligned}$$

For $\tilde{\boldsymbol\tau} \in T_N$, define $\boldsymbol\tau = (p^\gamma(\tilde\tau_1 - 1) + 1, \ldots, p^\gamma(\tilde\tau_s - 1) + 1)$, which implies $\left(\boldsymbol\tau, \tilde{\boldsymbol\beta}\right) \in \Gamma(p, N, s, r)$. Then the measurements become

$$\begin{aligned}
y_j &= \frac{1}{\sqrt p}\sum_{l=1}^s e^{-\frac{2\pi i(\tau_l - 1)(j-1)}{p}} \tilde\beta_l + z_j \\
&= \frac{1}{\sqrt p}\sum_{l=1}^s e^{-\frac{2\pi i p^\gamma(\tilde\tau_l - 1)(j-1)}{p}} \sqrt{\frac{rp\log p}{N}} + z_j \\
&= \frac{1}{\sqrt N}\sum_{l=1}^s e^{-\frac{2\pi i(\tilde\tau_l - 1)(j-1)}{N}} \sqrt{r\log p} + z_j
\end{aligned}$$

for $j = 1, \ldots, N$.

Denote by $\boldsymbol{F}_N$ the $N \times N$ normalized DFT matrix:

$$\boldsymbol{F}_N(j,k) = \frac{1}{\sqrt{N}} e^{-\frac{2\pi i(k-1)(j-1)}{N}}.$$

Define $\boldsymbol{\theta}(\tilde{\boldsymbol{\tau}}) \in \mathbb{R}^N$, such that the $\tilde{\tau}_l$th component of $\boldsymbol{\theta}(\tilde{\boldsymbol{\tau}})$ is $\sqrt{2r \log p} = \sqrt{2r(1-\gamma)\log N}$ for $l = 1, \ldots, s$, while other components are zeros. Then $\boldsymbol{\theta}(\tilde{\boldsymbol{\tau}})$ is a sparse vector, whose sparsity is $s = p^{1-\alpha} = N^{1-\frac{\alpha-\gamma}{1-\gamma}}$. Moreover, all its entries are all $\sqrt{2r \log p} = \sqrt{2r(1-\gamma)\log N}$. Now the measurements can be written as

$$\boldsymbol{y} = \frac{1}{\sqrt{2}} \boldsymbol{F}_N \boldsymbol{\theta}(\tilde{\boldsymbol{\tau}}) + \boldsymbol{z},$$

which is equivalent to

$$\sqrt{2}\boldsymbol{F}_N^* \boldsymbol{y} = \boldsymbol{\theta}(\tilde{\boldsymbol{\tau}}) + \sqrt{2}\boldsymbol{F}_N^* \boldsymbol{z}.$$

Notice that $\boldsymbol{F}_N^* \boldsymbol{z} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}, \boldsymbol{0})$. Since $\boldsymbol{\theta}(\tilde{\boldsymbol{\tau}})$ is deterministic, real and positive and the imaginary and real parts of $\sqrt{2}\boldsymbol{F}_N^* \boldsymbol{y}$ are independent, we have the following equivalent measurement:

$$\boldsymbol{v} := \Re(\sqrt{2}\boldsymbol{F}_N^* \boldsymbol{y}) = \boldsymbol{\theta}(\tilde{\boldsymbol{\tau}}) + \boldsymbol{w},$$

where $\boldsymbol{w} \in \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_N)$. Now the detection problem becomes nearly the standard sparse mean detection studied in [S33]; aslo see [S22, S30]. The only difference is that here $\tilde{\boldsymbol{\tau}} \in T_N$ satisfies the separation condition, which is

$$\tilde{\tau}_{l+1} - \tilde{\tau}_l \geq \log^2 N \text{ for } l = 1, \ldots, s-1, \tilde{\tau}_1 + N - \tilde{\tau}_s \geq \log^2 N.$$

We now prove that this difference is actually negligible. Suppose $\tilde{\boldsymbol{\tau}}$ is uniformly distributed in T_N. It induces a mixed simple alternative

$$p_1(\boldsymbol{v}) = \frac{1}{|T_N|} \sum_{\tilde{\boldsymbol{\tau}} \in T_N} p_{\tilde{\boldsymbol{\tau}}}(\boldsymbol{v}).$$

To prove that

$$\mathbb{P}_0(H_0 \text{ is rejected}) + \max_{(\boldsymbol{\beta},\boldsymbol{\tau}) \in \Gamma(p,N,s,r)} \mathbb{P}_{(\boldsymbol{\beta},\boldsymbol{\tau})}(H_0 \text{ is accepted}) \to 1,$$

by the standard Hellinger distance argument, it suffices to prove the

$$\mathbb{E}_0 \sqrt{\frac{p_1}{p_2}(\boldsymbol{v})} \geq 1 - o(1).$$

Suppose S_N is the collection of all subsets of $\{1, \dots, N\}$ with cardinality s. By Lemma A. 8 in [S30], we know $\frac{|T_N|}{|S_N|} = 1 - o(1)$. Define

$$\tilde{p}_1(\boldsymbol{v}) = \frac{1}{|S_N|} \sum_{\tilde{\boldsymbol{\tau}} \in S_N} p_{\tilde{\boldsymbol{\tau}}}(\boldsymbol{v}),$$

and define

$$\delta(\boldsymbol{v}) = \frac{|S_N|}{|T_N|}\tilde{p}_1 - p_1 = \frac{1}{|T_N|} \sum_{\tilde{\boldsymbol{\tau}} \in (S_N - T_N)} p_{\tilde{\boldsymbol{\tau}}}(\boldsymbol{v})$$

If $\tilde{\boldsymbol{\tau}}$ is uniformly distributed in S_N, $\tilde{p}_1$ becomes the simple mixed alternative, and the detection problem becomes the standard sparse mean detection problem. Notice that

$$\begin{aligned}
\mathbb{E}_0 \sqrt{\frac{p_1}{p_0}} &= \mathbb{E}_0 \sqrt{\frac{\frac{|S_N|}{|T_N|}\tilde{p}_1 - \delta}{p_0}} \\
&\geq \mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} - \mathbb{E}_0 \sqrt{\frac{\delta}{p_0}} \\
&\geq \mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} - \sqrt{\mathbb{E}_0 \frac{\delta}{p_0}} \\
&= \mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} - \sqrt{\frac{1}{|T_N|} \sum_{\tilde{\tau} \in S_N - T_N} 1} \\
&= \mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} - \sqrt{\frac{|S_N - T_N|}{|T_N|}} \\
&\geq \mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} - o(1).
\end{aligned}$$

Therefore, it suffices to prove $\mathbb{E}_0 \sqrt{\frac{\tilde{p}_1}{p_0}} \geq 1 - o(1)$. The problem now becomes the standard sparse mean vector detection studied in [S33, S22, S30]. Since $s = N^{1-\frac{\alpha-\gamma}{1-\gamma}}$ and the common nonzero components of the mean vector: $\sqrt{2r(1-\gamma)\log N}$, it suffices to require

$$\begin{cases} r(1-\gamma) < \frac{\alpha-\gamma}{1-\gamma} & \text{if } \frac{1}{2} < \frac{\alpha-\gamma}{1-\gamma} \leq \frac{3}{4}, \\ r(1-\gamma) < \left(1 - \sqrt{1 - \frac{\alpha-\gamma}{1-\gamma}}\right)^2 & \text{if } \frac{3}{4} < \frac{\alpha-\gamma}{1-\gamma} < 1. \end{cases}$$

This is exactly $r < \rho^*_\gamma(\alpha)$, and the proof is completed. □