1. Mutual coherence (40 points)

Recall that for an arbitrary pair of orthonormal bases \( \Psi = [\psi_1, \ldots, \psi_n] \in \mathbb{R}^{n \times n} \) and \( \Phi = [\phi_1, \ldots, \phi_n] \in \mathbb{R}^{n \times n} \), the mutual coherence \( \mu(\Psi, \Phi) \) of these two bases is defined by

\[
\mu(\Psi, \Phi) = \max_{1 \leq i,j \leq n} |\psi_i^\top \phi_j| \quad (1)
\]

(a) Show that

\[
\frac{1}{\sqrt{n}} \leq \mu(\Psi, \Phi) \leq 1.
\]

(b) Let \( \Psi = I \), and suppose that \( \Phi = [\phi_{i,j}]_{1 \leq i,j \leq n} \) is a Gaussian random matrix such that the \( \phi_{i,j} \)'s are i.i.d. random variables drawn from \( \phi_{i,j} \sim \mathcal{N}(0, 1/n) \). Can you provide an upper estimate on \( \mu(\Psi, \Phi) \) as defined in (1)? Since \( \Phi \) is a random matrix, we expect your answer to be a function \( f(n) \) such that

\[
P\{\mu(\Psi, \Phi) > f(n)\} \to 0 \quad \text{as} \quad n \quad \text{scales}.
\]

Hint: to simplify analysis, you are allowed to use the crude approximation

\[
P\{|z| > \tau\} \approx \exp(-\tau^2/2)
\]

for large \( \tau > 0 \), where \( z \sim \mathcal{N}(0, 1) \).

(c) Set \( n = 100 \). Generate a random matrix \( \Phi \) as in Part (b), and compute \( \mu(I, \Phi) \). Report the empirical distribution (i.e. histogram) of \( \mu(I, \Phi) \) out of 1000 simulations. How does your simulation result compare to your estimate in Part (b)?

(d) We now generalize the mutual coherence measure to accommodate a more general set of vectors beyond two bases. Specifically, for any given matrix \( A = [a_1, \ldots, a_p] \in \mathbb{R}^{n \times p} \) obeying \( n \leq p \), define the mutual coherence of \( A \) as

\[
\mu(A) = \max_{1 \leq i,j \leq p, i \neq j} \left| \frac{a_i^\top a_j}{\|a_i\| \|a_j\|} \right|.
\]

Show that

\[
\mu(A) \geq \sqrt{\frac{p-n}{p-1}} \cdot \frac{1}{n}.
\]

This is a special case of the Welch bound.

Hint: you may want to use the following inequality: for any positive semidefinite \( M \in \mathbb{R}^{n \times n} \), \( \|M\|_F^2 \geq \frac{1}{n} \left( \sum_{i=1}^{n} \lambda_i(M) \right)^2 \).

2. Picket-fence signal (10 points)

Suppose that \( \sqrt{n} \) is an integer. Let \( x \in \mathbb{R}^n \) be a picket-fence signal with uniform spacing \( \sqrt{n} \) such that

\[
x_i = \begin{cases} 
1, & \text{if } \frac{i-1}{\sqrt{n}} \text{ is an integer,} \\
0, & \text{else,}
\end{cases} \quad i = 1, \ldots, n.
\]
Compute
\[ \|x\|_0 \cdot \|Fx\|_0 \quad \text{and} \quad \|x\|_0 + \|Fx\|_0, \]
where \( F \) is the Fourier matrix such that
\[ (F)_{k,l} = \frac{1}{\sqrt{n}} \exp \left( -i \frac{2\pi (k-1)(l-1)}{n} \right), \quad 1 \leq k, l \leq n. \]

How do they compare to the uncertainty principles we derive in class?

3. \( \ell_1 \) minimization (20 points)

Suppose that \( A \) is an \( n \times 2n \) dimensional matrix. Let \( x \in \mathbb{R}^{2n} \) be an unknown \( k \)-sparse vector, and \( y = Ax \) the observed system output. This problem is concerned with \( \ell_1 \) minimization (or basis pursuit) in recovering \( x \), i.e.
\[ \minimize_{z \in \mathbb{R}^{2n}} \|z\|_1 \quad \text{s.t.} \quad Az = y. \tag{2} \]

(a) An optimization problem is called a linear program (LP) if it has the form
\[ \minimize_z \quad c^\top z + d \quad \text{s.t.} \quad Gz \leq h \quad Az = b \]
where \( c, d, G, h, A, \) and \( b \) are known. Here, for any two vectors \( r \) and \( s \), we say \( r \leq s \) if \( r_i \leq s_i \) for all \( i \). Show that (2) can be converted to a linear program.

(b) Set \( n = 256 \), and let \( k \) range between 1 and 128. For each choice of \( k \), run 10 independent numerical experiments: in each experiment, generate \( A = [a_{i,j}]_{1 \leq i \leq n, 1 \leq j \leq 2n} \) as a random matrix such that the \( a_{i,j} \)'s are i.i.d. standard Gaussian random variables, generate \( x \in \mathbb{R}^{2n} \) as a random \( k \)-sparse signal (e.g. you may generate the support of \( x \) uniformly at random, with each non-zero entry drawn from the standard Gaussian distribution), and solve (2) with \( y = Ax \). An experiment is claimed successful if the solution \( z \) returned by (2) obeys \( \|x - z\|_2 \leq 0.001 \|x\|_2 \). Report the empirical success rates (averaged over 10 experiments) for each choice of \( k \).