APPENDIX A

UNDER LS RECONSTRUCTION, $\Delta_1 \leq 0$

For LS we have:

$$R_S = U_k(M_S U_k)^\dagger.$$  

Lemma 1. For any matrix $A$, $\|U_k A\|_F^2 = \|A\|_F^2$.

Proof. Using Lemma 1,

$$\|U_k A\|_F^2 = \|U_k A A^T U_k^T\|_F^2 = \|A A^T\|_F^2 = \|A\|_F^2.$$  

Lemma 2. For LS, $\xi_1(S) = k - \operatorname{rank}(M_S U_k)$.

Proof. Using Lemma [1]

$$\xi_1(S) = \|I_k - \Pi\|_F^2$$

$$= \|I_k - \Pi(I_k - \Pi)^T\|_F^2$$

$$= \|I_k - \Pi(I_k - \Pi)^T\|_F^2$$

$$= \|I_k - \Pi - \Pi I_k\|_F^2$$

$$= \|I_k - \Pi\|_F^2$$

$$= k - \operatorname{rank}(M_S U_k).$$  

Lemma 3. For LS, $\Delta_1(S, v) \in \{0, -1\}$.

Proof. Removing a vertex from $S$ removes a row from $M_S U_k$, reducing the rank by 0 or 1.

$$\Delta_1(S, v) = \xi_1(S) - \xi_1(S \setminus \{v\})$$

$$= -\operatorname{rank}(M_S U_k) + \operatorname{rank}(M_S U_k)$$

$$\in \{0, -1\}.$$  

Non-positivity of $\Delta_1$ immediately follows from Lemma [3].

APPENDIX B

UNDER LS RECONSTRUCTION, $\Delta_1 < 0 \iff \Delta_2 > 0$

We first need the following lemmas.

Lemma 4.

$$\xi_2(S) = \frac{1}{\lambda_i^2} \sum \frac{1}{\lambda_i^2}$$  

where $\lambda_i^2$ is the $i$th eigenvalue of $(M_S U_k)(M_S U_k)^T$.

Proof. By definition and Appendix A Lemma [1]

$$\xi_2(S) = \frac{||R_S||_F^2}{||U_k (M_S U_k)^\dagger||_F^2}$$

$$= \frac{||U_k (M_S U_k)^\dagger||_F^2}{||M_S U_k||_F^2}$$

which is the sum of the squares of the singular values of $(M_S U_k)^\dagger$. The pseudoinverse maps the singular values of $M_S U_k$ onto the singular values of $(M_S U_k)^\dagger$ in the following way [23 Section 5.5.2]:

$$\sigma_i((M_S U_k)^\dagger) = \begin{cases} 0 & \text{if } \sigma_i(M_S U_k) = 0 \\ \sigma_i(M_S U_k)^{-1} & \text{otherwise} \end{cases}$$  

and the squares of the singular values of $M_S U_k$ are $\lambda_i^2$ [23 Eq. (8.6.1)]. Summing them gives the result.

Lemma 5.

$$\operatorname{rank}((M_S U_k)(M_S U_k)^T) = \operatorname{rank}(M_S U_k) \leq k.$$

Proof. For the equality: $\operatorname{rank}(M_S U_k)$ is the number of strictly positive singular values it has [23 Corollary 2.4.6]. By [23 Eq. (8.6.2)], this is the same as the number of strictly positive eigenvalues of $(M_S U_k)(M_S U_k)^T$, which is $\operatorname{rank}((M_S U_k)(M_S U_k)^T)$. 

For the inequality: $M_S U_k$ has $k$ columns and so must have column rank less than or equal to $k$. Row rank being equal to column rank gives the result.

Lemma 6. For LS, $\Delta_1 = 0 \iff \Delta_2 \leq 0$.

Proof. Note that $(M_{S \setminus \{v\}} U_k)(M_{S \setminus \{v\}} U_k)^T$ is a principal submatrix of $(M_S U_k)(M_S U_k)^T$. Write the eigenvalues of $(M_{S \setminus \{v\}} U_k)(M_{S \setminus \{v\}} U_k)^T$ as $\lambda_1, \ldots, \lambda_n$ and the eigenvalues of $(M_S U_k)(M_S U_k)^T$ as $\mu_1, \ldots, \mu_{n+1}$. Then by Cauchy’s Interlacing Theorem [24 p. 59],

$$0 \leq \mu_1 \leq \lambda_1 \leq \cdots \leq \lambda_n \leq \mu_{n+1} \leq 1$$  

where the outer bounds come from the fact that both matrices are principal submatrices of $U_k U_k^T$, an orthogonal projection matrix.

1) $\Delta_1 = 0 \implies \Delta_2 \leq 0$: $\Delta_1 = 0$ implies the rank of $M_S U_k$ does not change with the removal of $v$, so neither does the rank of $(M_S U_k)(M_S U_k)^T$. As the rank is unchanged, $(M_S U_k)(M_S U_k)^T$ has one more zero-eigenvalue than $(M_{S \setminus \{v\}} U_k)(M_{S \setminus \{v\}} U_k)^T$. This means:

$$\mu_1 = 0$$  

$$\lambda_i = 0 \iff \mu_{i+1} = 0$$  

By Cauchy’s Interlacing Theorem, $\lambda_i \leq \mu_{i+1}$ and so

$$\frac{1}{\lambda_i} \geq \frac{1}{\mu_{i+1}} \text{ if } \lambda_i \neq 0 \text{ and } \mu_{i+1} \neq 0.$$  

Therefore

$$\sum \frac{1}{\lambda_i^2} \geq \sum \frac{1}{\mu_{i+1}^2}$$  

as we have the same number of non-zero terms in each of these terms by [23 and 24], and the inequality is proved by...
summing over the non-zero terms using (25). Equation (26) is exactly
\[ \xi_2(S \setminus \{v\}) \geq \xi_2(S). \] (27)

Rearranging gives \( \Delta_2 \leq 0 \).

2) \( \Delta_1 = 0 \iff \Delta_2 \leq 0 \): We prove the equivalent statement
\[ \Delta_1 \neq 0 \implies \Delta_2 > 0. \] (28)

By Lemma 3, if \( \Delta_1 \neq 0 \) then \( \Delta_1 = -1 \). This means that the rank of \( M_S U_k \) is reduced by 1 by the removal of \( v \), therefore \( (M_S U_k)(M_S U_k)^T \) has one more non-zero eigenvalue than \( (M_{S \setminus \{v\}} U_k)(M_{S \setminus \{v\}} U_k)^T \). This means:
\[ \mu_{n+1} > 0 \] (29)
\[ \lambda_i \neq 0 \iff \mu_i \neq 0 \] (30)

By Cauchy’s interlacing theorem, \( \lambda_i \geq \mu_i \) and so
\[ \frac{1}{\lambda_i} \leq \frac{1}{\mu_i} \text{ if } \lambda_i \neq 0 \text{ and } \mu_i \neq 0. \] (31)

Let \( I \) be the number of zero eigenvalues of \( (M_S U_k)(M_S U_k)^T \). Then
\[ \sum_{I \leq i \leq n} \frac{1}{\lambda_i^2} \leq \sum_{I \leq i \leq n} \frac{1}{\mu_i^2} < \sum_{I \leq i \leq n+1} \frac{1}{\mu_i^2}. \] (32)

With the left inequality by matching terms via (30) and then summing over (31), and the right inequality because (29) means \( \frac{1}{\mu_i^2} > 0 \). We then note the left and the right terms in this equality say:
\[ \sum_{\lambda_i^2 \neq 0} \frac{1}{\lambda_i^2} < \sum_{\mu_i^2 \neq 0} \frac{1}{\mu_i^2} \] (33)

or equivalently,
\[ \xi_2(S \setminus \{v\}) < \xi_2(S). \] (34)

Rearranging gives \( \Delta_2 > 0 \).

We finally have the following:

**Lemma 7.** For LS, \( \Delta_1 < 0 \iff \Delta_2 > 0 \).

**Proof.** By Lemma 3 and Lemma 6.

**APPENDIX C**

**Proof of Theorem 1**

**Proof.** For brevity, we fix \( S \) and \( v \) and write \( \Delta_1 = \Delta_1(S, v) \) and \( \Delta_2 = \Delta_2(S, v) \).

Rearranging (14) gives us that \( v \) improves \( S \) if and only if
\[ \Delta_1 + \sigma^2 \cdot \Delta_2 > 0 \] (35)
or equivalently if and only if
\[ \Delta_1 > -\sigma^2 \cdot \Delta_2. \] (36)

By definition, \( \sigma^2 = \frac{k}{N \cdot \text{SNR}} \), so this condition is equivalent to
\[ \Delta_1 > -\frac{k}{N \cdot \text{SNR}} \Delta_2 \] (37)
and as SNR is strictly positive, this is equivalent to
\[ \text{SNR} \cdot \Delta_1 > -\frac{k}{N} \Delta_2 \] (38)

We can now use the major lemmas from the previous appendices. By Lemma 3 we have two possible values of \( \Delta_1(S, v) \):

\[ \Delta_1 = 0: \]

Lemma 6 means \( \Delta_2 < 0 \), so
\[ \Delta_1 + \sigma^2 \cdot \Delta_2 = \sigma^2 \cdot \Delta_2 < 0 \] (39)
and so \( v \) does not improve \( S \).

\[ \Delta_1 = -1: \]

Eq. (38) simplifies to:
\[ -\text{SNR} > -\frac{k}{N} \Delta_2 \] (40)
which is equivalent to
\[ \text{SNR} < \frac{k}{N} \Delta_2. \] (41)

On the one hand, \( v \) improves \( S \) implies \( \Delta_1 = -1 \), which implies (41). On the other hand, (41) implies \( \Delta_2 > 0 \) which in turn implies \( \Delta_1 = -1 \), which means (41) implies (38), which implies \( v \) improves \( S \).

Noting that the right-hand side of (41) is \( \tau(S, v) \), which completes the proof. □

**APPENDIX D**

**Proof of Theorem 2**

We restate the theorem:

**Theorem 4.** Consider any sequence of vertices \( v_1, \ldots, v_N \) with no repeated vertices, and let \( S_i = \{v_1, \ldots, v_i\} \). Then there are exactly \( k \) indices \( I_1, \ldots, I_k \) such that under LS reconstruction of a noisy \( k \)-bandlimited signal,
\[ \forall 1 \leq j \leq k : \tau(S_{I_j}, v_{I_j}) > 0 \] (42)
and so for some SNR > 0 removing \( v_{I_j} \) would improve \( S_{I_j} \).

**Proof.** By Appendix C Lemma 2,
\[ \xi_1(S_i) = k - \text{rank}(M_S U_k). \] (43)

By Appendix C Lemma 3 \( \Delta_1 \in \{0, -1\} \) and as \( \text{rank}(U_k) = k, \xi_1(S_N) = 0 \). As \( \xi_1(S_0) = k \), we must have exactly \( k \) indices for which \( \Delta_1(S_i, v_i) = -1 \), and by Appendix C Lemma 6 we have exactly \( k \) indices for which \( \Delta_2(S_i, v_i) > 0 \). As \( \tau(S_i, v_i) = \frac{k}{N} \Delta_2(S_i, v_i) \), we’re done. □
APPENDIX E
PROOF OF THEOREM 3

Proof. By Appendix C, Lemma 2, the noiseless error
\[ \xi_1(S) = k - \text{rank}(M_S U_k) \]  
(44)
must be 0, as we can perfectly reconstruct any \( k \)-bandlimited signal. Therefore, \( \text{rank}(M_S U_k) = k \).

\( M_S U_k \) is a \( k \times k \) matrix of full rank, so its rows must be
linearly independent. Any subset of linearly independent rows
is linearly independent, so for any non-empty \( R \subset S \), \( M_R U_k \) has linearly independent rows.

Greedy schemes pick increasing sample sets: that is, if asked
to pick a vertex sample set \( S_m \) of size \( m \) for \( m < k \) and a
sample set \( S \) of size \( k \), \( S_m \subset S \). Therefore for any sample
set \( S_m \) of size \( m \leq k \) picked by the scheme, \( M_{S_m} U_k \) has
independent rows.

If \( M_{S_m} U_k \) has independent rows, then removal of any row
(corresponding to removing any vertex) reduces its rank by 1; that is,
\[ \forall m \leq k : \forall v \in S_m : \Delta_1(S_m, v) = -1 \]  
(45)
Then, by Appendix C, Lemma 7,
\[ \forall m \leq k : \forall v \in S_m : \Delta_2(S_m, v) > 0 \]  
(46)
and as \( \tau(S_m, v) = \frac{k}{m} \Delta_2(S_m, v) \) and \( \frac{k}{m} > 0 \),
\[ \forall m \leq k : \forall v \in S_m : \tau(S_m, v) > 0 \]  
(47)
This proves (18).

As \( M_{S_k} U_k \) has \( k \) independent rows, it is of rank \( k \).
Adding further rows can’t decrease its rank, so for \( m' > k \),
\( \text{rank}(M_{S_{m'}} U_k) \geq k \). As \( U_k \) is of rank \( k \), \( \text{rank}(M_{S_{m'}} U_k) \leq k \). This means for all samples sizes \( m' > k \), \( \text{rank}(M_{S_{m'}} U_k) = k \). This says that further additions of rows do not change rank;
that is,
\[ \forall m' > k : \forall v \in S_{m'} \setminus S_k : \Delta_1(S_{m'}, v) = 0 \]  
(48)
Then, by Appendix C, Lemma 6,
\[ \forall m' > k : \forall v \in S_{m'} \setminus S_k : \Delta_2(S_{m'}, v) \leq 0 \]  
(49)
and, like for (18) as \( \tau(S_{m'}, v) = \frac{k}{m'} \Delta_2(S_{m'}, v) \) and \( \frac{k}{m'} > 0 \),
\[ \forall m' > k : \forall v \in S_{m'} \setminus S_k : \tau(S_{m'}, v) \leq 0 \]  
(50)
This proves (19).

APPENDIX F
PROOF OF REMARK 4

A-Optimality

A-optimality depends on the existence of the inverse of
\( (M_S U_k)(M_S U_k)^T \) existing, which requires it to be of full
rank. By Appendix C, Lemma 5, if an A-optimal scheme picks
a set \( S \) of size \( k \), then \( \text{rank}(M_S U_k) = k \). Therefore, \( S \) is
a uniqueness set \([17]\) and can perfectly reconstruct any \( k \)-
bandlimited signal.

D- and E-optimality

We show that for sample sizes less than \( k \) we can always
pick a row which keeps \( (M_S U_k)(M_S U_k)^T \) full rank (of rank
\( |S| \)), and that D- and E-optimal schemes do so.

By Appendix C, Lemma 5, \( \text{rank}(M_S U_k)(M_S U_k)^T = \text{rank}(M_S U_k) \), so we only need to ensure \( \text{rank}(M_S U_k) = |S| \).

We proceed by induction: given \( S_1 \) with \( |S_1| = 1 \),
\( \text{rank}(M_S U_k) = 1 \). Assume that for \( S_i \) with \( |S_i| = i < k \),
\( \text{rank}(M_S U_k) = i \). As \( \text{rank}(U_k) = k \) and \( i < k \), we can find
a row to add to \( M_S U_k \) which will increase its rank (else all
other rows would lie in the \( i \)-dimensional space spanned by
the rows of \( M_S U_k \), which would imply \( \text{rank}(U_k) = i \), which is
a contradiction as \( i < k \)). Adding the vertex which corresponds
to the row to \( S_i \) gives \( S_{i+1} \) with \( \text{rank}(M_{S_{i+1}} U_k) = i + 1 \).

We have shown that we can greedily choose to keep
\( \text{rank}(M_S U_k) = |S| \). We now show that D- and E-optimal
schemes do so. The eigenvalues of \( (M_S U_k)(M_S U_k)^T \) are
non-negative (see Appendix C, Eq. (22)), so any invertible
\( (M_S U_k)(M_S U_k)^T \) will have a strictly positive determinant
and minimum eigenvalue, which are preferable under the D-
and E-optimality criterion respectively to a non-invertible
\( (M_S U_k)(M_S U_k)^T \), which has a determinant and minimum
eigenvalue of 0. Therefore, greedy D- and E-optimal sampling
schemes will make sure \( (M_S U_k)(M_S U_k)^T \) is invertible,
and thus keep \( \text{rank}(M_S U_k) = |S| \) for \( |S| \leq k \). Therefore
when D- and E-optimal schemes pick a set \( S \) of size \( k \),
\( \text{rank}(M_S U_k) = k \). Therefore, \( S \) is a uniqueness set \([17]\)
and can perfectly reconstruct any \( k \)-bandlimited signal.

APPENDIX G
ADDITIONAL RESULTS

We show thresholds for the ER, BA and SBM graphs with
100 vertices (Fig. 5). We also present MSE plots for the larger
BA (Fig. 4) and SBM (Fig. 3) graphs.
Fig. 3: $\tau$ for different random graph models under LS reconstruction ($\#$vertices = 100, bandwidth = 10)

Fig. 4: Average MSE for LS reconstruction on BA Graphs ($\#$vertices=1000, bandwidth = 100) with different SNRs

Fig. 5: Average MSE for LS reconstruction on SBM Graphs ($\#$vertices=1000, bandwidth = 100) with different SNRs