APPENDIX A
COMBINATORIAL MATRIX ANALYSIS

Here, we prove Lem. 1, which is the main ingredient in our proof of Thm. 1. We then outline how additionally assuming the spark condition for B simplifies the proof and also allows for its extension to the case where only an upper bound on the number of columns m of A is known. This extension is applied to the proof of Thm. 3 in Appendix B.

We first prove some auxiliary lemmas. Given a collection of sets $\mathcal{T}$, let $\cap\mathcal{T}$ denote their intersection.

Lemma 3. *Let $M \in \mathbb{R}^{n\times m}$. If every $2k$ columns of M are linearly independent, then for any $\mathcal{T} \subseteq \bigcup_{\ell\le k}\binom{[m]}{\ell}$, we have:*

$$\text{Span}\{M_{\cap\mathcal{T}}\} = \bigcap_{S\in\mathcal{T}} \text{Span}\{M_S\}.$$

Proof. By induction, it is enough to prove the lemma when $|\mathcal{T}| = 2$. The proof now follows directly from the assumption. □

Lemma 4. *Fix $k \geq 2$. Let $\mathcal{V} = \{V_1, \ldots, V_k\}$ be subspaces of $\mathbb{R}^m$ and let $V = \bigcap \mathcal{V}$. For every $\mathbf{x} \in \mathbb{R}^m$, we have:*

$$|\mathbf{x} - \Pi_V \mathbf{x}|_2 \le \frac{1}{1-\xi(\mathcal{V})} \sum_{i=1}^{k} |x - \Pi_{V_i} x|_2, \tag{19}$$

provided $\xi(\mathcal{V}) \neq 1$, where ξ is given in Def. 3.

Proof. Fix $\mathbf{x} \in \mathbb{R}^m$ and $k \geq 2$. The proof consists of two parts. First, we shall show that:

$$|\mathbf{x}-\Pi_V\mathbf{x}|_2 \le \sum_{\ell=1}^{k} |\mathbf{x}-\Pi_{V_\ell}\mathbf{x}|_2 + |\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}\mathbf{x}-\Pi_V\mathbf{x}|_2. \tag{20}$$

For each $\ell \in \{2, \ldots, k+1\}$ (when $\ell = k+1$, the product $\Pi_{V_k}\cdots\Pi_{V_\ell}$ is set to I), we have by the triangle inequality and the fact that $\|\Pi_{V_\ell}\|_2 \le 1$ (as Π_{V_ℓ} are projections):

$$|\Pi_{V_k}\cdots\Pi_{V_\ell}\mathbf{x}-\Pi_V\mathbf{x}| \le |\Pi_{V_k}\cdots\Pi_{V_{\ell-1}}\mathbf{x}-\Pi_V\mathbf{x}| + |\mathbf{x}-\Pi_{V_{\ell-1}}\mathbf{x}|.$$

Summing these inequalities over ℓ gives (20).

Next, we show how the result (19) follows from (20) and the following result of [13, Thm. 9.33]:

$$|\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}\mathbf{x} - \Pi_V\mathbf{x}|_2 \le z|\mathbf{x}|_2 \quad \text{for} \quad \mathbf{x}\in\mathbb{R}^m, \tag{21}$$

for $z^2 = 1 - \prod_{\ell=1}^{k-1}(1-z_\ell^2)$ and $z_\ell = \cos\theta_F\left(V_\ell, \cap_{s=\ell+1}^k V_s\right)$. To see this, note that:

$$\begin{aligned} &|\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}(\mathbf{x}-\Pi_V\mathbf{x}) - \Pi_V(\mathbf{x}-\Pi_V\mathbf{x})|_2 && (22)\\ &\qquad = |\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}\mathbf{x} - \Pi_V\mathbf{x}|_2, && (23)\end{aligned}$$

since $\Pi_{V_\ell}\Pi_V = \Pi_V$ for all $\ell = 1, \ldots, k$ and $\Pi_V^2 = \Pi_V$. Therefore by (21) and (22), it follows that:

$$\begin{aligned} &|\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}\mathbf{x} - \Pi_V\mathbf{x}|_2 \\ &\quad = |\Pi_{V_k}\Pi_{V_{k-1}}\cdots\Pi_{V_1}(\mathbf{x}-\Pi_V\mathbf{x}) - \Pi_V(\mathbf{x}-\Pi_V\mathbf{x})|_2 \\ &\quad \le z|\mathbf{x}-\Pi_V\mathbf{x}|_2.\end{aligned}$$

Combining this with (20) and rearranging, we arrive at:

$$|\mathbf{x}-\Pi_V\mathbf{x}|_2 \le \frac{1}{1-z}\sum_{i=1}^{k}|\mathbf{x}-\Pi_{V_i}\mathbf{x}|_2. \tag{24}$$

Finally, since the ordering of the subspaces is arbitrary, we can replace z in (24) with $\xi(\mathcal{V})$ to obtain (19). □

Lemma 5. *Fix integers $k < m$, and let $T = \{S_1, \ldots, S_m\}$ be the set of contiguous length k intervals in some cyclic order of $[m]$. Suppose there exists a map $\pi : T \to \binom{[m]}{k}$ such that:*

$$|\bigcap_{i\in J}\pi(S_i)| \le |\bigcap_{i\in J} S_i| \quad \textit{for} \quad J \in \binom{[m]}{k}. \tag{25}$$

Then, $|\pi(S_{j_1}) \cap \cdots \cap \pi(S_{j_k})| = 1$ for all consecutive (modulo m) indices $j_1, \ldots, j_k$.

Proof. Consider the set $Q_m = \{(r,t) : r \in \pi(S_t), t \in [m]\}$, which has mk elements. By the pigeon-hole principle, there is some $q \in [m]$ and $J \in \binom{[m]}{k}$ such that $(q,j) \in Q_m$ for all $j \in J$. In particular, we have $q \in \cap_{j\in J}\pi(S_j)$ so that from (25) there must be some $p \in [m]$ with $p \in \cap_{j\in J}S_j$. Since $|J| = k$, this is only possible if the elements of $J = \{j_1, \ldots, j_k\}$ are consecutive modulo m, in which case $|\cap_{j\in J} S_j| = 1$. Hence, $|\cap_{j\in J}\pi(S_j)| = 1$ as well.

We next consider if any other $t \notin J$ is such that $q \in \pi(S_t)$. Suppose there were such a t; then, we would have $q \in \pi(S_t) \cap \pi(S_{j_1}) \cap \cdots \cap \pi(S_{j_k})$ and (25) would imply that the intersection of every k-element subset of $\{S_t\} \cup \{S_j : j \in J\}$ is nonempty. This would only be possible if $\{t\} \cup J = [m]$, in which case the result then trivially holds since then $q \in \pi(S_j)$ for all $j \in [m]$. Suppose now there exists no such t; then, letting $Q_{m-1} \subset Q_m$ be the set of elements of Q_m not having q as a first coordinate, we have $|Q_{m-1}| = (m-1)k$.

By iterating the above arguments, we arrive at a partitioning of Q_m into sets $R_i = Q_i \setminus Q_{i-1}$ for $i = 1, \ldots, m$, each having a unique element of $[m]$ as a first coordinate common to all k elements while having second coordinates which form a consecutive set modulo m. In fact, every set of k consecutive integers modulo m is the set of second coordinates of some R_i. This must be the case because for every consecutive set J we have $|\cap_{j\in J} S_j| = 1$, whereas if J is the set of second coordinates for two distinct sets R_i, we would have $|\cap_{j\in J}\pi(S_j)| \geq 2$, violating (25). □

Proof of Lem. 1 (Main Lemma). We assume $k \geq 2$ since the case $k = 1$ was proved at the beginning of Sec. III. Let $S_1, \ldots, S_m$ be length k intervals in some cyclic ordering of $[m]$. We begin by showing that $\dim(\text{Span}\{B_{\pi(S_i)}\}) = k$ for all $i \in [m]$. Fix $i \in [m]$ and note that by (8), all unit vectors $\mathbf{u} \in \text{Span}\{A_{S_i}\}$ satisfy $d(u, \text{Span}\{B_{\pi(S_i)}\}) \le \frac{\phi_k(A)}{\rho k}\delta$ for $\delta < \frac{L_2(A)}{\sqrt{2}}$. By definition of $L_2(A)$, for all 2-sparse $\mathbf{x} \in \mathbb{R}^m$:

$$L_2(A) \le \frac{|A\mathbf{x}|_2}{|\mathbf{x}|_2} \le \rho\frac{|\mathbf{x}|_1}{|\mathbf{x}|_2} \le \rho\sqrt{2}.$$

Hence, $\delta < \rho$. Since $\phi_k \le 1$, we have $d(u, \text{Span}\{B_{\pi(S_i)}\}) < 1$, and it follows by Lem. 2 that $\dim(\text{Span}\{B_{\pi(S_i)}\}) \geq \dim(\text{Span}\{A_{S_i}\}) = k$. Since $|\pi(S_i)| = k$, we in fact have $\dim(\text{Span}\{B_{\pi(S_i)}\}) = k$.

We will now show that:

$$|\bigcap_{i \in J} \pi(S_i)| \leq |\bigcap_{i \in J} S_i| \quad \text{for} \quad J \in \binom{[m]}{k}. \tag{26}$$

Fix $J \in \binom{[m]}{k}$. By (8), we have for all unit vectors $\mathbf{u} \in \cap_{i \in J}\text{Span}\{B_{\pi(S_i)}\}$ that $d(\mathbf{u}, A_{S_i}) \leq \frac{\phi_k(A)}{\rho k}\delta$ for all $j \in J$, where $\delta < \frac{L_2(A)}{\sqrt{2}}$. It follows from Lem. 4 that:

$$d\left(\mathbf{u}, \bigcap_{i \in J} \text{Span}\{A_{S_j}\}\right) \leq \frac{\delta}{\rho}\left(\frac{\phi_k(A)}{1 - \xi(\{A_{S_i} : i \in J\})}\right) \leq \frac{\delta}{\rho},$$

where the 2nd inequality follows from the definition of $\phi_k(A)$.

Now, since $\text{Span}\{B_{\cap_{i \in J}\pi(S_i)}\} \subseteq \cap_{i \in J}\text{Span}\{B_{\pi(S_i)}\}$ and (by Lem. 3) $\cap_{i \in J}\text{Span}\{A_{S_i}\} = \text{Span}\{A_{\cap_{i \in J}S_i}\}$, we have:

$$d\left(\mathbf{u}, A_{\cap_{i \in J}S_i}\right) \leq \frac{\delta}{\rho}, \quad \text{for unit } \mathbf{u} \in \text{Span}\{B_{\cap_{i \in J}\pi(S_i)}\}. \tag{27}$$

In particular, by Lem. 2 (since $\delta/\rho < 1$) we have that $\dim(\text{Span}\{B_{\cap_{i \in J}\pi(S_i)}\}) \leq \dim(\text{Span}\{A_{\cap_{i \in J}S_i}\})$ and (26) follows by the linear independence of the columns of A_{S_i} and $B_{\pi(S_i)}$ for all $i \in [m]$.

Suppose now that $J = \{i - k + 1, \ldots, i\}$ so that $\cap_{i \in J}S_i = \{i\}$. By (26), we have that $\cap_{i \in J}\pi(S_i)$ is either empty or it contains a single element. Lem. 5 ensures that the latter case is the only possibility. Thus, the association $i \mapsto \cap_{i \in J}\pi(S_i)$ defines a map $\hat{\pi} : [m] \to [m]$. Recalling (10), it follows from (27) that for all unit vectors $\mathbf{u} \in \text{Span}\{A_i\}$, we have $d\left(\mathbf{u}, B_{\hat{\pi}(i)}\right) \leq \delta/\rho$ also. Since i is arbitrary, it follows that for every basis vector $\mathbf{e}_i \in \mathbb{R}^m$, letting $c_i = |A\mathbf{e}_i|_2^{-1}$ and $\varepsilon = \delta/\rho$, there exists some $c_i' \in \mathbb{R}$ such that $|c_iA\mathbf{e}_i - c_i'B\mathbf{e}_{\hat{\pi}(i)}|_2 \leq \varepsilon$ where $\varepsilon < \frac{L_2(A)}{\sqrt{2}}\min_{j \in [m]} c_i$. This is exactly the supposition in (13) and the result follows from the subsequent arguments of Sec. III. ☐

The arguments above can easily be modified to prove the following variation of Lem. 1, key to proving Thm. 3.

Lemma 6 (Main Lemma for $m < m'$)**.** *Fix positive integers n, m, m', and k where $k < m < m'$, and let T be the set of intervals of length k in some cyclic ordering of $[m]$. Let $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{n \times m'}$ both satisfy spark condition* (2) *with A having maximum column ℓ_2-norm ρ. If there exists a map $\pi : T \to \binom{[m']}{k}$ and some $\delta < \frac{L_2(A)}{\sqrt{2}}$ such that for $S \in T$:*

$$\Theta(A_S, B_{\pi(S)}) \leq \frac{\delta}{\rho k}\min(\phi_k(A), \phi_k(B)), \tag{28}$$

then (9) *holds for some $n \times m$ submatrix of B.*

We state the required modifications briefly. Since $m' > m$, we may not invoke Lem. 5 (which requires $m = m'$) to show that $|\cap_{i \in J}\pi(S_i)| = 1$ for $J = \{i - k + 1, \ldots, i\}$. Instead, under the additional assumption that B satisfies the spark condition, we may simply swap the roles of A and B in the proof of (27) to show that $\dim(\text{Span}\{B_{\cap_{i \in J}\pi(S_i)}\}) = \dim(\text{Span}\{A_{\cap_{i \in J}S_i}\})$ and from which the required fact then follows. The map $\hat{\pi}$ is then defined similarly, only now with codomain $[m']$, thereby reducing the proof to the $k = 1$ case where the $n \times m$ submatrix of B is formed from the columns indexed by the image of $\hat{\pi}$.

Appendix B
Proofs of Thms. 2 & 3 and Cors. 1 & 2

Proof of Cor. 1. We need only demonstrate how to produce N vectors $\mathbf{a}_i$ such that for every interval of length k in some cyclic order on $[m]$, there are $(k-1)\binom{m}{k} + 1$ vectors in general linear position supported there. Let $\gamma_1, \ldots, \gamma_N$ be any distinct numbers. Then the columns of the $k \times N$ matrix $V = (\gamma_j^i)_{i,j=1}^{k,N}$ are in general linear position (since the γ_j are distinct, any $k \times k$ "Vandermonde" sub-determinant is nonzero). Next, fix a cyclic order on $[m]$ and let T be the set of contiguous length k intervals in the order. Finally, form the k-sparse vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N \in \mathbb{R}^m$ with supports $S \in T$ (partitioning the a_i evenly among these supports so that each contains $(k-1)\binom{m}{k} + 1$ vectors a_i) by setting the nonzero values $\mathbf{a}_i$ to be those contained in the ith column of V. ☐

We now determine classes of datasets Y having a stable sparse coding that are cut out by a single polynomial equation.

Proof of Thm. 2. We sketch the argument, leaving the details to the reader. Let M be the $n \times m$ matrix with columns $A\mathbf{a}_i$, $i \in [N]$. Consider the following polynomial [9, Sec. IV] in the entries of A and the $\mathbf{a}_i$:

$$g(A, \{\mathbf{a}_i\}_{i=1}^N) = \prod_{S \in \binom{[n]}{k}} \sum_{S' \in \binom{[N]}{k}} (\det M_{S',S})^2,$$

with notation as in Sec. II.

It can be checked that when g is nonzero for a substitution of real numbers for the indeterminates, all of the genericity requirements on A and $\mathbf{a}_i$ in our proofs of stability in Thm. 1 are satisfied (in particular, the spark condition on A). The statement of the theorem now follows directly. ☐

Proof of Cor. 2. First, note that if a set of measure spaces $\{(X_i, \Sigma_i, \nu_i)\}_{i=1}^p$ is such that ν_i is absolutely continuous with respect to μ for all $i = 1, \ldots, p$, where μ is the standard Borel measure on $\mathbb{R}$, then the product measure $\prod_{i=1}^p \nu_i$ is absolutely continuous with respect to the standard Borel product measure on $\mathbb{R}^p$. By Thm. 2, there is a polynomial such that $A \in \mathbb{R}^{n \times m}$ and $\mathbf{a}_i$ which is nonzero whenever Y has a stable k-sparse representation in $\mathbb{R}^m$; in particular, this property (stability) holds with probability one. ☐

Proof of Thm. 3. The proof is very similar to the proof of Thm. 1 in Sec. III, the difference being that now we establish a map $\pi : [m] \to [m']$ satisfying the requirements of Lem. 6 by pigeonholing $(k-1)\binom{m'}{k} + 1$ vectors with respect to holes $[m']$ and eventually applying Lem. 6 in place of Lem. 1. The value of C in this case is:

$$C = \left(\frac{\sqrt{k^3}}{\min(\phi_k(A), \phi_k(B))}\right)\frac{\max_{j \in [m]}|A_j|_2}{\min_{S \in T} L_k(AX_{I(S)})}. \tag{29}$$

The same manipulations in Rem. 3 show how (5) then follows from (4) for some subset of m coefficients in the $\mathbf{b}_i$. ☐