Supplementary Material

7.1. Proof of Proposition 1

Proof. 1. Any symmetric tensor Q that satisfies the conditions in part 1 of Proposition 1 is dual feasible. The decomposition measure $\mu^\star$ is primal feasible. We also have

$$\begin{aligned}\langle Q, A\rangle &= \sum_{p=1}^{r} \lambda_p \langle Q, x^p \otimes x^p \otimes x^p\rangle \\ &= \sum_{p=1}^{r} \lambda_p q(x^p) = \sum_{p=1}^{r} \lambda_p = \mu^\star(\mathbb{S}^{n-1}),\end{aligned}$$

establishing a zero duality gap at the primal-dual feasible solution pair $(\mu^\star, Q)$. Therefore, $\mu^\star$ is primal optimal and Q is dual optimal.

For uniqueness, suppose $\hat{\mu}$ is another optimal solution. We then have

$$\begin{aligned}\mu^\star(\mathbb{S}^{n-1}) &= \langle Q, A\rangle \\ &= \left\langle Q, \int_{\mathbb{S}^{n-1}} x \otimes x \otimes x d\hat{\mu} \right\rangle \\ &= \sum_{x \in \operatorname{supp}(\mu^\star)} \hat{\mu}(x) q(x) \\ &+ \int_{\mathbb{S}^{n-1}/\operatorname{supp}(\mu^\star)} q(x) d\hat{\mu} \\ &< \sum_{x_p \in \operatorname{supp}(\mu^\star)} \hat{\lambda}_p + \int_{\mathbb{S}^{n-1}/\operatorname{supp}(\mu^\star)} 1 d\hat{\mu} \\ &= \hat{\mu}(\mathbb{S}^{n-1})\end{aligned}$$

due to condition (14) if $\hat{\mu}(\mathbb{S}^{n-1}/\operatorname{supp}(\mu^\star)) > 0$, contradicting the optimality of $\hat{\mu}$. So all optimal solutions are supported on $\operatorname{supp}(\mu^\star)$. Since the tensors $\{x^p \otimes x^p \otimes x^p, p = 1, \dots, r\}$ are linearly independent, the coefficients are also uniquely determined.

2. Denote by p_0 and d_0 the optimal values for the primal problem (4) and the dual problem (5), respectively; and denote by p_1 and d_1 the optimal values for the primal-dual problems (9) and (12) (or (10)), respectively. We next argue that these four quantities are equal under the conditions in part 2. First, part 1 establishes $p_0 = d_0$. Second, weak duality and the construction of relaxations (9) and (12) imply that $d_1 \leq p_1 \leq p_0 = d_0$. Also the feasible set of (12) projected onto the Q space is a subset of the feasible set of (5). Since the conditions of part 2 state that the optimal dual solution Q of (5) is also feasible to (12), we hence conclude that Q is also an optimal solution of (12) and obtain $d_1 = d_0$. Therefore, $p_0 = d_0 = d_1 = p_1$, and the relaxations (9) and (12) are tight.

To show the optimality of $y^\star$, the $2k$-truncation of the (infinite) moment vector $\bar{y}^\star$ corresponding to the measure $\mu^\star$. We first note that $y^\star$ is feasible to (9). Then zero duality gap, as verified below

$$y_0^\star = \mu^\star(\mathbb{S}^{n-1}) = p_0 = d_1 = \langle Q, A\rangle,$$

establishes the optimality of $y^\star$.

3. Denote by $\sigma(x) = \nu_k(x)' H \nu_k(x)$ the SOS polynomial associated with H. Note $\nu_k(x^p)' H \nu_k(x^p) = \sigma(x^p) = 1 - q(x^p) = 0$ for $p = 1, \dots, r$, implying $H\nu_k(x^p) = 0, p = 1, \dots, r$ due to $H \succcurlyeq 0$. Since $\operatorname{rank}(H) = |\mathbb{N}_k^n| - r$ by the assumption, the null space of H is $\operatorname{span}\{\nu_k(x^p), p = 1, \dots, r\}$.

For any optimal solution $\hat{y}$ of (9), complementary slackness implies that

$$H M_k(\hat{y})) = 0.$$

So the eigen-space corresponding to the non-zero eigenvalues of $M_k(\hat{y})$ is a subspace of $\operatorname{span}\{\nu_k(x^p), p = 1, \dots, r\}$. We hence write

$$M_k(\hat{y}) = VDV'$$

where $V = \left[\nu_k(x^1) \cdots \nu_k(x^r)\right]$ and D is an $r \times r$ semidefinite matrix (not necessarily diagonal at this point). Note that $M_k(y^\star) = V \Lambda V'$ where $\Lambda = \operatorname{diag}([\lambda_1, \dots, \lambda_r])$. We next argue that $D = \Lambda$.

The moment matrix $M_k(\hat{y})$ contains a known submatrix specified by the third order moments in the tensor A, and hence is equal to the corresponding submatrix in $M_k(y^\star)$. More precisely, $M_k(\hat{y})$ contains the block (at the location indicated by the orange color in Figure 5):

$$\begin{aligned}&\int_{\mathbb{S}^{n-1}} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \begin{bmatrix} x_1^2 & x_1 x_2 & \cdots & x_{n-1} x_n & x_n^2 \end{bmatrix} d\mu^\star \\ =& X \Lambda V_2'\end{aligned}$$

where $X = \left[x^1 \cdots x^p\right]$, and V_2 is the submatrix of V whose rows correspond to the second-order monomials in $\nu_k(x)$. Therefore, we have

$$X \Lambda V_2' = X D V_2' \tag{25}$$

According to Lemma 3.1 (ii) of (De Lathauwer, 2008), $\operatorname{rank}(X) = r$ implies $\operatorname{rank}(V_2) = r$. Multiplying both sides of (25) by the pseudo-inverse matrices $X^\dagger$ from the left and $(V_2')^\dagger$ from the right yield $D = \Lambda$. So $M_k(\hat{y}) = M_k(y^\star)$, and $\hat{y} = y^\star$ is the unique solution of (9).

To see that we can extract the measure $\mu^\star$ from $M_k(\hat{y}) = M_k(y^\star)$, we note that the matrix $M_k(y^\star) = V \Lambda V'$ has rank r for all $k \geq 1$. Hence the flat extension condition $\operatorname{rank}(M_{k-1}(y^\star) = M_k(y^\star))$ is satisfied for all $k \geq 2$. Therefore, according to (Curto & Fialkow, 1996; Henrion & Lasserre, 2005), we could recover the measure from the moment matrix $M_k(y^\star)$. □

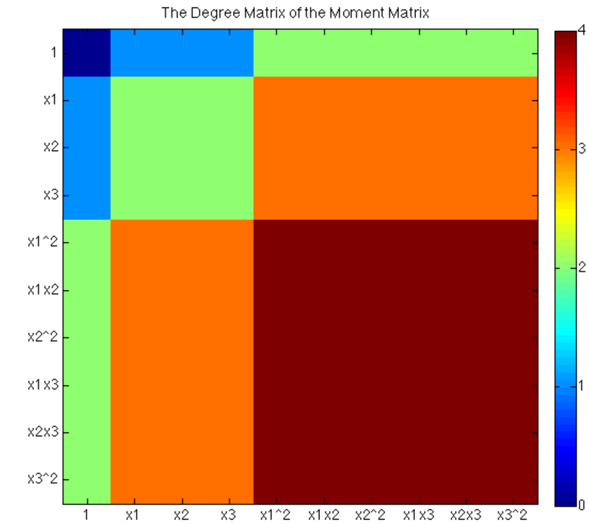

Figure 5. The colors encode the degrees of the entries in the moment matrix for an instance with $n = 3, k = 2$.

7.2. Dual Certificate: the Orthonormal Case

The proof of Theorem 1 is based on a perturbation analysis of the orthogonal case, which is the focus of this and the next sections. Hereafter, the relaxation order is fixed to $k = 2$.

When the vectors $\{x^p, p = 1, \ldots, r\}$ are orthonormal, we verify that the symmetric tensor

$$Q = \sum_{p=1}^{r} x^p \otimes x^p \otimes x^p$$

satisfies the conditions in part 1 of Proposition 1. To see this, note

$$q(x^p) = \langle Q, x^p \otimes x^p \otimes x^p \rangle = \sum_{p'=1}^{r} \langle x^{p'}, x^p \rangle^3 = 1.$$

Moreover, for any fixed $x \in \mathbb{S}^{n-1}$ we have

$$\begin{aligned} q(x) = \langle Q, x \otimes x \otimes x \rangle &= \sum_{p=1}^{r} \langle x^p, x \rangle^3 \\ &\leq \max_p \langle x^p, x \rangle \sum_{p=1}^{r} \langle x^p, x \rangle^2 \\ &\leq \|X^T x\|_2^2 \end{aligned}$$

where we used $\max_p \langle x^p, x \rangle \leq \max_p \|x^p\| \|x\| = 1$ for all p. Due to the orthogonality of the columns of $X = [x^1 \cdots x^r]$, we clearly have $\|X^T x\|_2^2 \leq \|x\|_2^2 = 1$. For $q(x) = 1$, all the involved inequalities must be equalities. For $\max_p \langle x^p, x \rangle = 1$, we need $x = x^p$ for some p, which is the only possible case that $q(x) = 1$. For all other cases, $q(x) < 1$. Therefore, $Q = \sum_p x^p \otimes x^p \otimes x^p$ satisfies the conditions of part 1 in Proposition 1. This argument combined with part 1 of Proposition 1 lead to

Theorem 3. *If the vectors in* $\operatorname{supp}(\mu^\star)$ *are orthonormal, then* $\mu^\star$ *is the unique optimal solution to* (4).

7.3. SOS Dual Certificate: the Orthonormal Case

In the following, we show that for $q(x) = \sum_{p=1}^{r} \langle x, x^p \rangle^3$, we can find an SOS $\sigma(x)$ and a polynomial $s(x)$ with degrees 4 and 2 respectively, such that

$$1 - q(x) = \sigma(x) + s(x)(\|x\|_2^2 - 1).$$

We start with assuming $x^p = e_p$, the pth canonical basis vector, for $p = 1, 2, \ldots, r$, in which case $q(x)$ becomes $\sum_{p=1}^{r} x_p^3$. Later on we will apply a rotation to derive the general case from this special case.

We set

$$s(x) = -\frac{3}{2}\left(\sum_{p=1}^{r} x_p^2\right) - \frac{3}{2}\left(\sum_{p=r+1}^{n} x_p^2\right) = \nu_1(x)' G_0 \nu_1(x)$$

where

$$G_0 := \begin{bmatrix} 0 & \\ & -\frac{3}{2} I_n \end{bmatrix}. \tag{26}$$

Consider

$$\begin{aligned} & 1 - q(x) - s(x)(\|x\|_2^2 - 1) \\ =& 1 - \sum_{p=1}^{r} x_p^3 + \frac{3}{2}\left(\sum_{p=1}^{r} x_p^2\right)\left(\sum_{p=1}^{n} x_p^2 - 1\right) \\ & + \frac{3}{2}\left(\sum_{p=r+1}^{n} x_p^2\right)\left(\sum_{p=1}^{n} x_p^2 - 1\right) \\ =& 1 - \frac{3}{2}\left(\sum_{p=1}^{r} x_p^2\right) - \frac{3}{2}\left(\sum_{p=r+1}^{n} x_p^2\right) - \sum_{p=1}^{r} x_p^3 \\ & + \frac{3}{2}\sum_{p=1}^{r} x_p^4 + \frac{3}{2}\sum_{p=r+1}^{n} x_p^4 \\ & + 3\sum_{p<p'=1}^{r} x_p^2 x_{p'}^2 + 3\sum_{p<p'=r+1}^{n} x_p^2 x_{p'}^2 + 3\sum_{p=1}^{r}\sum_{p'=1}^{n} x_p^2 x_{p'}^2. \end{aligned} \tag{27}$$

We show that this polynomial is an SOS $\sigma(x)$ with Gram matrix H_0 defined on top of the next page. Here the row partition of H_0 corresponds to the partition of the Veronese

$$H_0 := \begin{bmatrix} 1 & & & & & & -\mathbf{1}_r' & f\mathbf{1}_{n-r}' \\ & \frac{1}{2}I_r & & & & & -\frac{1}{2}I_r & \\ & & aI_{n-r} & & & & & \\ & & & I_{C_2^r} & & & & \\ & & & & bI_{r(n-r)} & & & \\ & & & & & cI_{C_2^{n-r}} & & \\ -\mathbf{1}_r & -\frac{1}{2}I_r & & & & & \frac{1}{2}I_r + \mathbf{1}_r\mathbf{1}_r' & d\mathbf{1}_r\mathbf{1}_{n-r}' \\ f\mathbf{1}_{n-r} & & & & & & d\mathbf{1}_{n-r}\mathbf{1}_r' & \left(\frac{3}{2}-e\right)I_{n-r} + e\mathbf{1}_{n-r}\mathbf{1}_{n-r}' \end{bmatrix} \quad (28)$$

map $\nu_2(x)$ given in the following

$$\nu_2(x) \quad := \begin{bmatrix} \nu_2^0(x) \\ \nu_2^1(x) \\ \nu_2^2(x) \\ \nu_2^3(x) \\ \nu_2^4(x) \\ \nu_2^5(x) \\ \nu_2^6(x) \\ \nu_2^7(x) \end{bmatrix} \quad (29)$$

with

$$\nu_2^0(x) = 1$$
$$\nu_2^1(x) = \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix}$$
$$\nu_2^2(x) = \begin{bmatrix} x_{r+1} \\ \vdots \\ x_n \end{bmatrix}$$
$$\nu_2^3(x) = \begin{bmatrix} x_1x_2 \\ x_1x_3 \\ \vdots \\ x_{r-1}x_r \end{bmatrix}$$
$$\nu_2^4(x) = \begin{bmatrix} x_1x_{r+1} \\ \vdots \\ x_rx_n \end{bmatrix}$$
$$\nu_2^5(x) = \begin{bmatrix} x_{r+1}x_{r+2} \\ \vdots \\ x_{n-1}x_n \end{bmatrix}$$
$$\nu_2^6(x) = \begin{bmatrix} x_1^2 \\ \vdots \\ x_r^2 \end{bmatrix}$$
$$\nu_2^7(x) = \begin{bmatrix} x_{r+1}^2 \\ \vdots \\ x_n^2 \end{bmatrix}$$

and a, b, c, d, e, f are parameters to be determined later.

Since

$$\begin{aligned} &\nu_2(x)'H_0\nu_2(x) \\ =&1 - \frac{3}{2}\sum_{p=1}^r x_p^2 + (a+2f)\sum_{p=r+1}^n x_p^2 - \sum_{p=1}^r x_p^3 \\ &+ \frac{3}{2}\sum_{p=1}^r x_p^4 + \frac{3}{2}\sum_{p=r+1}^n x_p^4 \\ &+ 3\sum_{p<p'=1}^r x_p^2x_{p'}^2 + (c+2e)\sum_{p<p'=r+1}^n x_p^2x_{p'}^2 \\ &+ (b+2d)\sum_{p=1}^r\sum_{p'=1}^n x_p^2x_{p'}^2 \end{aligned}$$

comparison of coefficients with those of $1 - q(x) - s(x)(\|x\|_2^2 - 1)$ in (27) gives

$$\begin{aligned} a + 2f &= -\frac{3}{2} \\ c + 2e &= 3 \\ b + 2d &= 3 \end{aligned}$$

We will judiciously choose the parameters so that H_0 is PSD. Note that H_0 must have r zero eigenvalues with eigenvectors $\{\nu_2(e^p) : p = 1, \dots, r\}$. For later analysis, we also need H_0 to have precisely r zero eigenvalues, and the smallest non-zero eigenvalue of H_0 to be lower bounded by a numerical constant regardless of n and r.

For that purpose, we next find all the eigenvalues of H_0. The obvious ones include $a, 1, b$ and c of multiplicities $n-r, C_2^r, r(n-r)$ and C_2^{n-r}, respectively. The rest of eigenvalues are those of E defined as

$$\begin{bmatrix} 1 & & -\mathbf{1}_r' & f\mathbf{1}_{n-r}' \\ & \frac{1}{2}I_r & -\frac{1}{2}I_r & \\ -\mathbf{1}_r & -\frac{1}{2}I_r & \frac{1}{2}I_r + \mathbf{1}_r\mathbf{1}_r' & d\mathbf{1}_r\mathbf{1}_{n-r}' \\ f\mathbf{1}_{n-r} & & d\mathbf{1}_{n-r}\mathbf{1}_r' & \left(\frac{3}{2}-e\right)I_{n-r} + e\mathbf{1}_{n-r}\mathbf{1}_{n-r}' \end{bmatrix}$$

We choose $e + a = \frac{3}{2}$ and decompose E as $A + B$ such that A is

$$\begin{bmatrix} 1 & & -\mathbf{1}_r' & f\mathbf{1}_{n-r}' \\ & \frac{1}{2r}\mathbf{1}_r\mathbf{1}_r' & -\frac{1}{2r}\mathbf{1}_r\mathbf{1}_r' & \\ -\mathbf{1}_r & -\frac{1}{2r}\mathbf{1}_r\mathbf{1}_r' & \left(1+\frac{1}{2r}\right)\mathbf{1}_r\mathbf{1}_r' & d\mathbf{1}_r\mathbf{1}_{n-r}' \\ f\mathbf{1}_{n-r} & & d\mathbf{1}_{n-r}\mathbf{1}_r' & (e + \frac{a}{(n-r)})\mathbf{1}_{n-r}\mathbf{1}_{n-r}' \end{bmatrix}$$

and B is

$$\begin{bmatrix} 0 & & & \\ & \frac{1}{2}\left(I_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r'\right) & -\frac{1}{2}\left(I_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r'\right) & \\ & -\frac{1}{2}\left(I_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r'\right) & \frac{1}{2}\left(I_r - \frac{1}{r}\mathbf{1}_r\mathbf{1}_r'\right) & \\ & & & * \end{bmatrix}$$

where the bottom-right block of B occupied by $*$ is $a\left(I_{n-r} - \frac{1}{n-r}\mathbf{1}_{n-r}\mathbf{1}_{n-r}'\right)$. It is easy to verify that $AB = BA = 0$. Hence the eigenvalues of E consist of those of A and B. The eigenvalues of B are 0, 1, and a of multiplicities $r+3, r-1, n-r-1$, respectively.

Next we choose the parameters such that the eigenvalues of A are easy to compute. We first ensure that A has rank 3, which, by rank invariance of Gaussian elimination, requires the following matrix,

$$\begin{bmatrix} 1 & & & \\ & \frac{1}{2r}\mathbf{1}_r\mathbf{1}_r' & & \\ & & \mathbf{0}_r & (d+f)\mathbf{1}_r\mathbf{1}_{n-r}' \\ & & (d+f)\mathbf{1}_{n-r}\mathbf{1}_r' & * \end{bmatrix}$$

whose bottom-right block is $\left(e + \frac{a}{(n-r)} - f^2\right)\mathbf{1}_{n-r}\mathbf{1}_{n-r}'$, to have rank 3, or equivalently, $d + f = 0$.

Multiplying A with a vector of the form $v := \begin{bmatrix} \alpha \\ \beta\mathbf{1}_r \\ \gamma\mathbf{1}_r \\ \delta\mathbf{1}_{n-r} \end{bmatrix}$ shows that the eigenvectors of A are of the form v. Consequently, the non-zero eigenvalues of A can be computed by solving a smaller set of eigenvalue equations

$$\begin{bmatrix} 1 & 0 & -r & f(n-r) \\ 0 & 1/2 & -1/2 & 0 \\ -1 & -1/2 & r+1/2 & -f(n-r) \\ f & 0 & -fr & (n-r)e+a \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = \lambda \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \tag{30}$$

We already have five equations on a, b, c, d, e, f:

$$\begin{aligned} a + 2f &= -\frac{3}{2} \\ c + 2e &= 3 \\ b + 2d &= 3 \\ e + a &= \frac{3}{2} \\ d + f &= 0 \end{aligned}$$

or,

$$\begin{aligned} b &= 3 - 2d = 3 - \frac{3}{2} - a = \frac{3}{2} - a \\ c &= 3 - 2e = 2a \\ d &= \frac{3}{4} + \frac{a}{2} \\ e &= \frac{3}{2} - a \\ f &= -\frac{3}{4} - \frac{a}{2} \end{aligned}$$

Plugging these into the matrix in (30) leads to a matrix involving a single parameter a:

$$\begin{bmatrix} 1 & 0 & -r & -\left(\frac{3}{4}+\frac{a}{2}\right)(n-r) \\ 0 & 1/2 & -1/2 & 0 \\ -1 & -1/2 & r+1/2 & \left(\frac{3}{4}+\frac{a}{2}\right)(n-r) \\ -\left(\frac{3}{4}+\frac{a}{2}\right) & 0 & \left(\frac{3}{4}+\frac{a}{2}\right)r & (n-r)\left(\frac{3}{2}-a\right)+a \end{bmatrix}$$

Symbolic calculation shows that the non-zero eigenvalues of this matrix are zeros of the polynomial

$$\begin{aligned} &h(\lambda; r, n, a) = (2+r)(15(-n+r) \\ &+ 4a(-4 + (7+a)n - (7+a)r)) \\ &+ 2(16 + 39n - 31r + 15(n-r)r \\ &- 4a^2(n-r)(1+r) + 4a((1+r)(8+7r) - n(11+7r)))\lambda \\ &+ 16(-4 - 3n + 2a(-1+n-r) + r)\lambda^2 + 32\lambda^3 \end{aligned}$$

We want to make sure $\lambda = a \neq 0$ is one non-zero eigenvalue, which means $h(a; r, n, a) = 0$, or after simplification:

$$\begin{aligned} &a^3(r-3) + 15(r+2) + 4a^2(13r+32) - 2a(29r+67) \\ &= 0 \end{aligned}$$

We pick the smallest positive root branch $a = a(r)$, which is an increasing function of r with limit $a(+\infty) = \frac{1}{2}$, and $a(1) > 0.3387$. We next argue that, after plugging $a = a(r)$, $h(\lambda; r, n, a(r))$ has two other zeros that are larger than $\frac{1}{2}$ (hence larger than $a(r)$), which means the other two non-zero eigenvalues of A are greater than $a(r) \in (0.3387, 0.5)$. The argument is based on median value theorem by showing $h(1/2; r, n, a(r)) > 0$, $h(n/2; r, n, a(r)) < 0$ combined with the obvious fact $\lim_{\lambda\to\infty} h(\lambda; r, n, a(r)) = +\infty$.

We first show $h(1/2; r, n, a) > 0$ for $1 \leq r \leq n$ and $a \in [0.2, 0.5)$. As a function of r with parameters n and a, the function

$$\begin{aligned} h(1/2; r, n, a) = 4 - 8a - 3n + 20an + 4a^2n \\ + (3 - 20a - 4a^2)r \end{aligned}$$

is linear in r and is decreasing since $3 - 20a - 4a^2 < 0$ for

$a \in [0.2, 0.5)$. Therefore, we obtain

$$\begin{aligned} h(1/2; r, n, a) &\geq h(1/2; n, n, a) \\ &= 4 - 8a \\ &> 0. \end{aligned}$$

Second, we show that $h(n/2; r, n, a) < 0$ for $a \in [0.2, 0.5)$ and $r \in [0, n]$:

$$\begin{aligned} &h(n/2; r, n, a) \\ =&(-2+n)(16a + (7-4a(9+a))n + 8(-1+a)n^2) \\ &+ (30 - 8a(9+a) - 46n + 8a(11+a)n \\ &+ (19 - 4a(9+a))n^2)r + (-1+2a)(15+2a)(-1+n)r^2 \\ \leq&(-2+n)(16a + (7-4a(9+a))n \\ &+ 8(-1+a)n^2) + 2(1-2a)(15+2a)(-1+n)nr \\ &+ (-1+2a)(15+2a)(-1+n)r^2. \end{aligned}$$

We used the fact that

$$\begin{aligned} &30 - 8a(9+a) - 46n + 8a(11+a)n + (19 - 4a(9+a))n^2 \\ &\leq 2(1-2a)(15+2a)(n-1)n \end{aligned}$$

which can be proved by observing that

$$\begin{aligned} &2(1-2a)(15+2a)(n-1)n - (30 - 8a(9+a) - 46n \\ &+ 8a(11+a)n + (19 - 4a(9+a))n^2) \\ &= -30 + 8a(9+a) \\ &+ (46 - 8a(11+a) - 2(1-2a)(15+2a))n \\ &+ (-19 + 4a(9+a) + 2(1-2a)(15+2a))n^2 \end{aligned}$$

is an increasing function of n (since $(46 - 8a(11+a) - 2(1-2a)(15+2a)) > 0$ for $a \in [0.2, 0.5)$), and its value at $n = 1$ is $-3 + 12a(9+a) - 8a(11+a) \geq 1$.

Now the upper bound on $h(n/2; r, n, a)$ is an increasing function of r for $r \in [1, n]$. We therefore further bound $h(n/2; r, n, a)$ by setting $r = n$ in its upper bound:

$$\begin{aligned} h(n/2; r, n, a) \leq& -32a - 14n + 8a(11+a)n \\ &+ 8(1-3a)n^2 + (7 - 4a(5+a))n^3 \\ :=& u(n; a) \end{aligned}$$

Since $\frac{d}{dn}u(n; a)$ is

$$-14 + 8a(11+a) + 16(1-3a)n + 3(7 - 4a(5+a))n^2,$$

which is decreasing for $n \geq 0$ due to $3(7 - 4a(5+a)) < 0$ and $16(1-3a) < 0$ when $a \in (0.3387, 0.5)$, we have

$$\begin{aligned} \frac{d}{dn}u(n; a) \leq& \frac{d}{dn}u(8; a) \\ =& 1458 - 8a(517 + 95a) \\ <& 0 \end{aligned}$$

for $n \geq 8$ and $a \in (0.3387, .5)$. Therefore, $u(n; a)$ is further upper bounded by its value at $n = 8$ for $n \geq 8$:

$$\begin{aligned} h(n/2; r, n, a) \leq& u(8; a) = -16(-249 + 2a(347 + 62a)) \\ <& 0 \end{aligned}$$

for $a \in (0.3387, .5)$.

To sum, we have showed that $h(\lambda; r, n, a(r))$, whose zeros are eigenvalues of A, has the property that $\lambda_1 = a(r) \in (0.3387, 1/2)$ is a zero, and $h(1/2; r, n, a(r)) > 0, h(n/2; r, n, a(r)) < 0$, and $h(+\infty; r, n, a) > 0$. Therefore, the other two zeros of $h(\lambda; r, n, a(r))$ are greater than $1/2 > a(r)$.

Therefore, the matrix H_0 has rank $|\mathbb{N}_2^n| - r$ and the minimal non-zero eigenvalue for H_0 is

$$\min\left\{a(r), \frac{3}{2} - a(r), 2a(r), \frac{1}{2}, 1\right\} = a(r)$$

since $a(r) \in (0.3387, 1/2)$. This shows that, when $\{x^p = e_p, p = 1, \ldots, r\}$, the matrix H_0 is PSD with rank $|\mathbb{N}_2^n| - r$ and the minimal non-zero eigenvalue is greater than $1/3$.

When $\text{supp}(\mu^\star)$ is orthonormal, but is not a subset of the canonical basis vectors, we augment the matrix $X = \begin{bmatrix} x^1 & \cdots & x^r \end{bmatrix}$ to an orthonormal matrix $P = \begin{bmatrix} X & P_1 \end{bmatrix}$ and transform the variable x to $z = P'x = P^{-1}x$. Then the tensor $A = \sum_p \lambda_p x^p \otimes x^p \otimes x^p$ is transformed to $\sum_p \lambda_p e_p \otimes e_p \otimes e_p$. So the dual polynomial

$$q_0(z) = 1 - \nu_2(z)' H_0 \nu_2(z) + \frac{3}{2}\|z\|_2^2(\|z\|_2^2 - 1)$$

with H_0 constructed above satisfies the conditions in Proposition 1, and certifies the optimality of the decomposition $\sum_p \lambda_p e_p \otimes e_p \otimes e_p$. We transform this polynomial back to the x-domain to obtain

$$\begin{aligned} q(x) :=& q_0(P'x) \\ =& 1 - \nu_2(P'x)' H_0 \nu_2(P'x) + \frac{3}{2}\|x\|_2^2(\|x\|_2^2 - 1) \end{aligned}$$

where we have used $\|P'x\|_2^2 = \|x\|_2^2$ since P is orthonormal. According to the change of basis formular in Lemma 1, the polynomial

$$\nu_2(P'x) H_0 \nu_2(P'x) = \nu_2(x)'(J' H_0 J)\nu_2(x)$$

is an SOS with the Gram matrix $J' H_0 J$, whose smallest eigenvalue is greater than $\frac{1}{2} \times \frac{1}{3} > \frac{1}{6}$. One can verify that $q(x)$ satisfies all the conditions in Proposition 1. As a consequence, we obtain:

Theorem 4. *If the vectors in* $\text{supp}(\mu^\star)$ *are orthonormal, then the SDP relaxation* (9) *with* $k = 2$ *gives the exact decomposition. Furthermore, the constructed dual polynomial has the form*

$$q(x) = 1 - \nu_2(x)' H \nu_2(x) + \frac{3}{2}\|x\|_2^2(\|x\|_2^2 - 1)$$

where H *has* r *zero eigenvalues, and the* $(r+1)$*th smallest eigenvalue is greater than* $\frac{1}{6}$*. When the support is formed* *by a subset of the canonical basis vectors, the lower bound on the $(r+1)$th smallest eigenvalue can be chosen as $\frac{1}{3}$.*

The SOS matrix decomposition is verified by Matlab. With $n = 7$ and $r = 3$, we have the following plot for H_0:

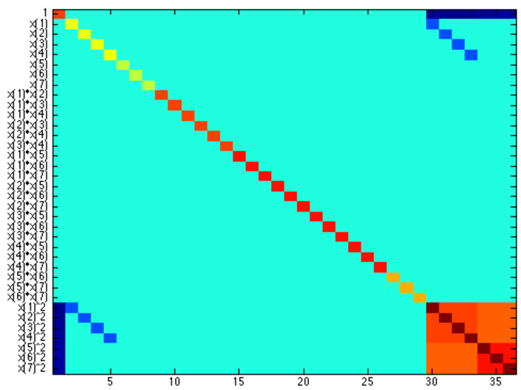

Figure 6. H_0 has $r = 4$ zero eigenvalues and the 5th smallest is $a(4) = 0.3789$.

7.4. Dual Certificate: The Non-Orthonormal Case

We now proceed to apply a perturbation analysis to construct a dual polynomial for the non-orthonormal case.

Suppose the measure $\mu^\star = \sum_{k=1}^r \lambda_k \delta(x - x^k)$ where $\{x^k, k = 1, \ldots, r\}$ are not orthogonal. Define $X = [x^1, \cdots, x^r]$ and find $P_1 \in \mathbb{R}^{n\times(n-r)}$ which has orthonormal columns and is orthogonal to X. Further define $P = \begin{bmatrix} X & P_1 \end{bmatrix}$. Then the transformation $x \mapsto z = P^{-1}x$ maps x^k to the kth canonical basis vector e_k. The unit sphere is mapped to an ellipsoid $E^{n-1} = \{z : z'P'Pz = 1\}$.

If we could construct a polynomial $q(z) = \langle Q, z \otimes z \otimes z\rangle$ with symmetric Q such that

$$q(e_k) = 1, k = 1, \ldots, r \tag{31}$$

$$|q(z)| < 1, z \in E^{n-1}, z \neq e_k \tag{32}$$

then the polynomial $q_1(x) := q(P^{-1}x) = \langle Q, P^{-1}x \otimes P^{-1}x \otimes P^{-1}x\rangle$ would satisfy

$$q_1(x^k) = q(e_k) = 1, k = 1, \ldots, r$$
$$|q_1(x)| = |q(P^{-1}x)| < 1, x \in \mathbb{S}^{n-1}, x \neq x^k.$$

The desired $q(z)$ must satisfy that $q(e_k) = 1$ and $q(z)$ achieves maximum at $z = e_k$ for $k = 1, \ldots, r$. Denote $L(z; \nu) = q(z) - \nu(z'P'Pz - 1)$ as the Lagrangian. A necessary condition for $q(z)$ to achieve maximum at e_k is given by the KKT condition:

$$\begin{aligned}\frac{\partial L(z)}{\partial z}|_{z=e_k} =& \frac{\partial q(z)}{\partial z}|_{z=e_k} - \nu \frac{\partial}{\partial z}(z'P'Pz - 1)|_{z=e_k} \\ =& 3\sum_{i=1}^n \langle Q, e_k \otimes e_k \otimes e_i\rangle e_i - 2\nu P'Pe_k \\ =& 0\end{aligned}$$

Taking inner product with e_k yields

$$3q(e_k) = 3\langle Q, e_k \otimes e_k \otimes e_k\rangle = 2\nu e_k' P'Pe_k = 3,$$

implying $\nu = \frac{3}{2}$. Therefore, the symmetric tensor Q must satisfy

$$\sum_{i=1}^n \langle Q, e_k \otimes e_k \otimes e_i\rangle e_i = P'Pe_k, k = 1, \ldots, r. \tag{33}$$

Note the condition (31) is a consequence of (33). We pick

$$\begin{aligned}Q =& \sum_{k=1}^r e_k \otimes e_k \otimes P'Pe_k + \sum_{k=1}^r e_k \otimes P'Pe_k \otimes e_k \\ &+ \sum_{k=1}^r P'Pe_k \otimes e_k \otimes e_k - 2\sum_{k=1}^r e_k \otimes e_k \otimes e_k\end{aligned}$$

which actually has minimal energy among all symmetric Qs that satisfy (33). The dual polynomial is then given by

$$\begin{aligned}q(z) =& \langle Q, z \otimes z \otimes z\rangle \\ =& \sum_{k=1}^r [3z_k^2(z'P'Pe_k) - 2z_k^3] \\ =& \sum_{k=1}^r [3(z'P'Pe_k) - 2z_k]z_k^2.\end{aligned}$$

Clearly, $q(z)$ satisfies the interpolation condition (31). In the following, we show that $q(z)$ also satisfies the condition (32). The argument is based on partitioning the ellipsoid E^{n-1} into a region that is far from any e_k and a region that is near to some e_k.

First note

$$q(z) \leq \max_k [3(z'P'Pe_k) - 2z_k] \sum_{k=1}^r z_k^2$$

When $z \in E^{n-1}$, due to $\|P'P - I\| \leq \epsilon$, we have $-\epsilon z'z \leq 1 - z'z \leq \epsilon z'z$, implying

$$\frac{1}{1+\epsilon} \leq \quad z'z \quad \leq \frac{1}{1-\epsilon}$$

Therefore, we can further upper bound $q(z)$ as

$$\begin{aligned}q(z) \leq& \max_k [3(z'P'Pe_k) - 2z_k] \sum_{k=1}^r z_k^2 \\ \leq& \frac{1}{1-\epsilon} \max_k [3(z'P'Pe_k) - 2z_k]\end{aligned}$$

So, if

$$\max_k[3(z'P'Pe_k) - 2z_k] < 1 - \epsilon$$

then $q(z) < 1$. Therefore, we have showed that $q(z) < 1$ in the "far-away" region.

Define $N_k = \{z : 3(z'P'Pe_k) - 2z_k \geq 1 - \epsilon, z'P'Pz = 1\}$. When $P'P \approx I$, this is saying $z_k \geq 1 - \epsilon$, so $z \in N_k$ is close to e_k. The union of N_ks defines the "near region".

We want to make sure that $q(z)$ is strictly less than 1 in each N_k except when $z = e_k \in N_k$. For that purpose, we perform a Taylor expansion of the Lagrangian $L(z) := L(z; 3/2)$ in N_k around $z = e_k$

$$\begin{aligned} L(z) &= q(z) - \frac{3}{2}(z'P'Pz - 1) \\ &= L(e_k) + (z - e_k)' \frac{\partial L}{\partial z}|_{z=e_k} \\ &+ \frac{1}{2}(z - e_k)'H(\xi_z)(z - e_k) \\ &= 1 + \frac{1}{2}(z - e_k)'H(\xi_z)(z - e_k) \end{aligned}$$

where $H(\xi_z)$ is the Hessian of $L(z)$ evaluated at ξ_z and $\xi_z \in L_{k,z} = \{tz + (1-t)e_k : t \in (0, 1)\}$, the line segment connecting e_k and z.

Since $q(z) = L(z)$ on the ellipsoid E^{n-1}, it suffices to show $\frac{1}{2}(z - e_k)'H(\xi_z)(z - e_k) < 0$ for $z \in N_k/\{e_k\}$. For this purpose, we compute the Hessian matrix $H(\xi)$:

$$\begin{aligned} H(\xi) &= \frac{\partial}{\partial z}\left[3\sum_{i=1}^{n}\langle Q, z \otimes z \otimes e_i\rangle e_i - 3P'Pz\right]|_{z=\xi} \\ &= 6\sum_{i,j=1}^{n}\langle Q, \xi \otimes e_j \otimes e_i\rangle e_i \otimes e_j - 3P'P \end{aligned}$$

Plugging in the expression of Q, we get that the Hessian $H(\xi)$ equals

$$\begin{aligned} &6\sum_{i,j=1}^{n}[\xi_j e_i'P'Pe_j + \xi_i e_j'P'Pe_i]e_i \otimes e_j \\ &\quad + 6\sum_{i=1}^{n}[(\xi'P'Pe_i) - 2\xi_i]e_i \otimes e_i - 3P'P \end{aligned}$$

To get a sense why this Hessian guarantees a negative second order term in the Taylor expansion, we set $\xi = e_k$ to get

$$\begin{aligned} H(e_k) &= 6\sum_{i,j=1}^{n}[e_k(j)e_i'P'Pe_j + e_k(i)e_j'P'Pe_i]e_i \otimes e_j \\ &+ 6\sum_{i=1}^{n}[(e_k'P'Pe_i) - 2e_k(i)]e_i \otimes e_i - 3P'P \\ &= 6\left[\sum_i(e_i'P'Pe_k)e_i \otimes e_k + \sum_j(e_j'P'Pe_k)e_k \otimes e_j\right] \\ &+ 6\sum_{i=1}^{n}[(e_k'P'Pe_i) - 2e_k(i)]e_i \otimes e_i - 3P'P \end{aligned}$$

When $P'P \approx I$,

$$\begin{aligned} H(e_k) &\approx 12e_k \otimes e_k - 6e_k \otimes e_k - 3I \\ &= 6e_k \otimes e_k - 3I \end{aligned}$$

which is PSD except in the direction e_k, which is orthogonal to the tangent space of $E^{n-1} \approx S^{n-1}$ at $z = e_k$. Therefore, the Hessian projected onto the tangent space is negative definite, as desired.

Returning to the non-orthogonal case, we bound

$$\begin{aligned} H(\xi) &= 6\sum_{i,j=1}^{n}[\xi_j e_i'P'Pe_j + \xi_i e_j'P'Pe_i]e_i \otimes e_j \\ &\quad + 6\sum_{i=1}^{n}[(\xi'P'Pe_i) - 2\xi_i]e_i \otimes e_i - 3P'P \end{aligned}$$

for $\xi \in L_{k,z}$ with $z \in N_k$, where

$N_k = \{z : 3(z'P'Pe_k) - 2z_k \geq 1 - \epsilon, z'P'Pz = 1\}$.

The simplifications

$$\begin{aligned} \sum_{i,j=1}^{n}(\xi_j e_i'P'Pe_j)e_i \otimes e_j &= P'P\,\mathrm{diag}(\xi) \\ \sum_{i,j=1}^{n}(\xi_i e_j'P'Pe_i)e_i \otimes e_j &= \mathrm{diag}(\xi)P'P \\ \sum_{i=1}^{n}(\xi'P'Pe_i)e_i \otimes e_i &= \mathrm{diag}(P'P\xi) \\ \sum_{i=1}^{n}\xi_i e_i \otimes e_i &= \mathrm{diag}(\xi) \end{aligned}$$

lead to the following compact expression for the Hessian matrix $H(\xi)$:

$$\begin{aligned} &6(P'P\,\mathrm{diag}(\xi) + \mathrm{diag}(\xi)P'P + \mathrm{diag}(P'P\xi) - 2\,\mathrm{diag}(\xi)) \\ &\quad - 3P'P \end{aligned}$$

We want to show that

$$(z - e_k)'H(\xi)(z - e_k) \quad < \quad 0, \forall\xi \in L_{k,z}, z \in N_k.$$

We first argue that $z \in N_k = \{z : 3(z'P'Pe_k) - 2z_k \geq 1 -$

$\epsilon, z'P'Pz = 1\}$ imposes certain restrictions on the size of z, and implies that z is close to e_k. Indeed, $\|I - P'P\| \le \epsilon$ and $z'P'Pz = 1$ imply that

$$\frac{1}{1+\epsilon} \le \frac{1}{\lambda_{\max}(P'P)} \le \|z\|_2^2 \le \frac{1}{\lambda_{\min}(P'P)} \le \frac{1}{1-\epsilon}.$$

To show the closeness of z and e_k, we observe that

$$\begin{aligned} 3z'P'Pe_k - 2z_k &= 3z'(P'P - I)e_k + 3z'e_k - 2z_k \\ &= z_k + 3z'(P'P - I)e_k \end{aligned}$$

Since $|3z'(P'P - I)e_k| \le 3\|z\|_2\|P'P - I\| \le \frac{3\epsilon}{\sqrt{1-\epsilon}}$, z_k is bounded from below as follows:

$$\begin{aligned} z_k &\ge 1 - \epsilon - 3z'(P'P - I)e_k \\ &\ge 1 - \epsilon - \frac{3\epsilon}{\sqrt{1-\epsilon}}. \end{aligned}$$

On the other hand, $z_k \le \|z\|_2 \le \frac{1}{\sqrt{1-\epsilon}}$.

A consequence of the sizes of z and z_k is that

$$\begin{aligned} \|z - z_k e_k\|_2^2 &= \sum_{j\neq k} z_j^2 \\ &= \|z\|_2^2 - z_k^2 \\ &\le \frac{1}{1-\epsilon} - \left(1 - \epsilon - \frac{3\epsilon}{\sqrt{1-\epsilon}}\right)^2 \end{aligned}$$

Therefore, we have

$$\begin{aligned} &\|z - e_k\|_\infty \\ \le& \max\{\epsilon + \frac{3\epsilon}{\sqrt{1-\epsilon}}, \frac{1}{\sqrt{1-\epsilon}} - 1, \\ &\sqrt{\frac{1}{1-\epsilon} - \left(1 - \epsilon - \frac{3\epsilon}{\sqrt{1-\epsilon}}\right)^2}\} \\ :=& c_1(\epsilon) \\ =& O(\epsilon) \end{aligned}$$

and

$$\begin{aligned} &\|z - e_k\|_2^2 \\ =& \sum_{j\neq k} z_j^2 + (z_k - 1)^2 \le \|z - z_k e_k\|_2^2 \\ &+ \max\left\{\epsilon + \frac{3\epsilon}{\sqrt{1-\epsilon}}, \frac{1}{\sqrt{1-\epsilon}} - 1\right\}^2 \\ =& \frac{1}{1-\epsilon} - \left(1 - \epsilon - \frac{3\epsilon}{\sqrt{1-\epsilon}}\right)^2 \\ &+ \max\left\{\epsilon + \frac{3\epsilon}{\sqrt{1-\epsilon}}, \frac{1}{\sqrt{1-\epsilon}} - 1\right\}^2 \\ =& c_2(\epsilon) \end{aligned}$$

Since $\xi_z \in L_{k,z}$, we have $\xi_z = tz + (1-t)e_k$ for some $t \in (0,1)$. As consequence, we obtain the following estimates for ξ_z:

$$\begin{aligned} \|\xi_z - e_k\|_\infty &\le t\|z - e_k\|_\infty \le c_1(\epsilon), \\ \|\xi_z - e_k\|_2^2 &\le t^2\|z - e_k\|_2^2 \le c_2(\epsilon), \\ \|\xi_z\|_2 &\le t\|z\|_2 + (1-t)\|e_k\|_2 \le \frac{1}{\sqrt{1-\epsilon}}. \end{aligned}$$

For notational simplicity, in the following we ignore the subscript z in ξ_z. We show that each term in

$$P'P\operatorname{diag}(\xi) + \operatorname{diag}(\xi)P'P + \operatorname{diag}(P'P\xi) - 2\operatorname{diag}(\xi)$$

is close to $e_k e_k'$, except the last term which is close to $2e_k e_k'$. The first term is bounded as follows:

$$\begin{aligned} &\|P'P\operatorname{diag}(\xi) - e_k e_k'\| \\ \le& \|P'P\operatorname{diag}(\xi) - P'Pe_k e_k'\| + \|P'Pe_k e_k' - e_k e_k'\| \\ \le& \|P'P\|\|\xi - e_k\|_\infty + \|P'P - I\| \\ \le& (1+\epsilon)c_1(\epsilon) + \epsilon \end{aligned}$$

Similar bounds hold for the term $\operatorname{diag}(\xi)P'P$:

$$\begin{aligned} &\|\operatorname{diag}(P'P\xi) - e_k e_k'\| \\ =& \|P'P\xi - e_k\|_\infty \\ \le& \|P'P\xi - \xi\|_\infty + \|\xi - e_k\|_\infty \\ \le& \|P'P - I\|\|\xi\|_2 + c_1(\epsilon) \\ \le& \frac{\epsilon}{\sqrt{1-\epsilon}} + c_1(\epsilon), \end{aligned}$$

and the term $\operatorname{diag}(\xi)$:

$$\|\operatorname{diag}(\xi) - e_k e_k'\| \le \|\xi - e_k\|_\infty \le c_1(\epsilon).$$

These bounds imply that

$$\begin{aligned} &\|P'P\operatorname{diag}(\xi) + \operatorname{diag}(\xi)P'P + \operatorname{diag}(P'P\xi) - 2\operatorname{diag}(\xi) \\ &\qquad - e_k e_k'\| \\ \le& 2(1+\epsilon)c_1(\epsilon) + 2\epsilon + \frac{\epsilon}{\sqrt{1-\epsilon}} + c_1(\epsilon) + c_1(\epsilon) \\ :=& c_3(\epsilon) \\ =& O(\epsilon). \end{aligned}$$

Furthermore, we have

$$\begin{aligned} &\|P'Pe_k e_k' P'P - e_k e_k\| \\ =& \|P'Pe_k e_k' P'P - P'Pe_k e_k' + P'Pe_k e_k' - e_k e_k'\| \\ \le& \|P'P\|\|e_k e_k'\|\|P'P - I\| + \|P'P - I\|\|e_k e_k'\| \\ \le& (1+\epsilon)\epsilon + \epsilon \\ =& O(\epsilon). \end{aligned}$$

Therefore, we get

$$\begin{aligned} &\|H(\xi) - (6P'Pe_k e_k' P'P - 3P'P)\| \\ \le& 6c_3(\epsilon) + 6\epsilon(2+\epsilon) \\ :=& c_4(\epsilon) \\ =& O(\epsilon). \end{aligned}$$

For any $z \in N_k$, we next show that $(z - e_k)'P'Pe_k$ is small due to the fact that both z and e_k lie on E^{n-1}:

$$\begin{aligned}1 &= z'P'Pz\\ &= e_k'P'Pe_k + 2(z-e_k)'P'Pe_k + (z-e_k)'P'P(z-e_k)\\ &= 1 + 2(z-e_k)'P'Pe_k + (z-e_k)'P'P(z-e_k)\end{aligned}$$

implying

$$\begin{aligned}|(z-e_k)'P'Pe_k| &= \frac{1}{2}(z-e_k)P'P(z-e_k)\\ &\le \frac{1}{2}\|P'P\|\|z-e_k\|_2^2\\ &\le \frac{1}{2}(1+\epsilon)\|z-e_k\|_2^2\end{aligned}$$

The following chain of inequalities

$$\begin{aligned}&(z-e_k)'H(\xi)(z-e_k)\\ \le&(z-e_k)'(6P'Pe_ke_k'P'P - 3P'P)(z-e_k)\\ &+\|z-e_k\|_2^2c_4(\epsilon)\\ =&6[(z-e_k)'P'Pe_k]^2 - 3(z-e_k)'P'P(z-e_k)\\ &+\|z-e_k\|_2^2c_4(\epsilon)\\ =&\frac{3}{2}(1+\epsilon)^2\|z-e_k\|_2^4 - 3(z-e_k)'P'P(z-e_k)\\ &+\|z-e_k\|_2^2c_4(\epsilon)\\ \le&\frac{3}{2}(1+\epsilon)^2\|z-e_k\|_2^4 - 3(1-\epsilon)\|z-e_k\|_2^2\\ &+\|z-e_k\|_2^2c_4(\epsilon)\\ =&\frac{3}{2}(1+\epsilon)^2\|z-e_k\|_2^4 - (3-3\epsilon-c_4(\epsilon))\|z-e_k\|_2^2\end{aligned}$$

show that the second order term is negative if

$$\frac{3}{2}(1+\epsilon)^2\|z-e_k\|_2^2 < 3-3\epsilon-c_4(\epsilon)$$

So it suffices to require

$$c_2(\epsilon)\frac{3}{2}(1+\epsilon)^2 < 3-3\epsilon-c_4(\epsilon)$$

Numerical computation shows that the above inequality holds if

$$\epsilon \leqslant 0.0016.$$

We summarize the above argument into a theorem:

Theorem 5. *For a symmetric tensor $A = \sum_{p=1}^r \lambda_p x^p \otimes x^p \otimes x^p$, if the vectors $\{x^p\}$ are near orthogonal, that is, the matrix $X = [x^1\ x^2\ \cdots\ x^r]$ satisfies*

$$\|X'X - I_r\| \le 0.0016,$$

then there exists a dual symmetric tensor Q such that the dual polynomial $q(x) = \langle Q, x\otimes x\otimes x\rangle$ satisfies the conditions in part 1 of Proposition 1. Thus, $A = \sum_{p=1}^r \lambda_p x^p \otimes x^p \otimes x^p$ is the unique decomposition that achieves the tensor nuclear norm, and can be found by solving (4).

7.5. SOS Dual Certificate: The Non-Orthonormal Case

After rotating to the canonical basis vectors, the dual polynomial we constructed for the orthogonal case is

$$q_0(z) = \sum_{k=1}^r z_k^3$$

while the one for the non-orthogonal case is

$$q(z) = \sum_{k=1}^r [3(z'P'Pe_k) - 2z_k]z_k^2.$$

We first show that $1-q(z)$ is an SOS modulo the ellipsoid E^{n-1}. We know that $q_0(z)$ is an SOS modulo the sphere, that is, there exist symmetric matrices $H \succcurlyeq 0$ and $G \in \mathbb{R}^{(n+1)\times(n+1)}$ such that

$$1-q_0(z) = \nu_2(z)'H\nu_2(z) + \nu_1(z)'G\nu_1(z)(\|z\|_2^2-1).$$

In Section 7.3, we constructed $G = G_0$ in (26) and $H = H_0$ in (28). So (H_0, G_0) is in the feasible set of the following two constraints:

$$\begin{aligned}&\nu_2(z)'H\nu_2(z) + \nu_1(z)'G\nu_1(z)(\|z\|_2^2-1) = 1-q_0(z), \forall z\\ &H \succcurlyeq 0.\end{aligned} \tag{34}$$

Note that any feasible H must satisfy $\nu_2(e_i)'H\nu_2(e_i) = 0$ for $i = 1, 2, \ldots, r$, implying that $\{\nu_2(e_i) : i = 1, 2, \ldots, r\}$ spans a subspace of the null space of H.

Define matrices B_α and C_α^0 that satisfy

$$\begin{aligned}\nu_2(z)\nu_2(z)' &= \sum_{|\alpha|\le 4} B_\alpha z^\alpha\\ \nu_1(z)\nu_1(z)'(\|z\|_2^2-1) &= \sum_{|\alpha|\le 4} C_\alpha^0 z^\alpha\end{aligned}$$

These notations allow us to write

$$\nu_2(z)'H\nu_2(z) = \langle \nu_2(z)\nu_2(z)', H\rangle = \sum_{|\alpha|\le 4}\langle B_\alpha, H\rangle z^\alpha$$

and

$$\begin{aligned}\nu_1(z)'G\nu_1(z)(\|z\|_2^2-1) &= \langle \nu_1(z)\nu_1(z)'(\|z\|_2^2-1), G\rangle\\ &= \sum_{|\alpha|\le 4}\langle C_\alpha^0, G\rangle z^\alpha\end{aligned}$$

Denote by b_α^0 the coefficient for z^α in $1-q_0(z)$. We write the polynomial equation $\nu_2(z)'H\nu_2(z) + \nu_1(z)'G\nu_1(z)(\|z\|_2^2-1) = 1-q_0(z)$ equivalently as

$$\langle B_\alpha, H\rangle + \langle C_\alpha^0, G\rangle = b_\alpha^0, |\alpha| \le 4$$

Therefore, we obtain the SDP formulation of (34)

$$\begin{aligned}&\text{find } G, H\\ &\text{subject to } \langle B_\alpha, H\rangle + \langle C_\alpha^0, G\rangle = b_\alpha^0, |\alpha|\le 4\\ &\qquad H \succcurlyeq 0.\end{aligned} \tag{35}$$

As aforementioned, G_0 and H_0 defined respectively in (26) and (28) form a feasible point for (35).

Now we switch to the non-orthogonal case, and we would like to show that

$$q(z) = \sum_{k=1}^{r} [3(z'P'Pe_k) - 2z_k] z_k^2$$

is an SOS module the ellipsoid E^{n-1}. That is, we want to solve the feasibility problem

$$\begin{aligned} &\text{find } G \text{ and } H \\ &\text{subject to} \\ &\nu_2(z)' H \nu_2(z) + \nu_1(z)' G \nu_1(z)(z'P'Pz - 1) = 1 - q(z) \\ &H \succcurlyeq 0. \end{aligned} \tag{36}$$

or equivalently in SDP

$$\begin{aligned} &\text{find } G \text{ and } H \\ &\text{subject to} \\ &\langle B_\alpha, H\rangle + \langle C_\alpha, G\rangle = b_\alpha, |\alpha| \le 4 \\ &H \succcurlyeq 0. \end{aligned} \tag{37}$$

Here B_α is defined as before, while b_α is the coefficient for z^α in $1 - q(z)$ for $|\alpha| \le 4$ and C_α is defined via

$$\nu_1(z)\nu_1(z)'(z'P'Pz - 1) = \sum_{|\alpha|\le 4} C_\alpha z^\alpha$$

We again note that any feasible H must satisfy $\nu_2(e_i)' H \nu_2(e_i) = 0$ for $i = 1, 2, \ldots, r$, implying that $\{\nu_2(e_i) : i = 1, 2, \ldots, r\}$ spans a subspace of the null space of H.

When $\|P'P - I\| \le \epsilon$ with ϵ small, C_α is close to C_α^0 and b_α is close to b_α^0. We claim that, when ϵ is sufficiently small, we can always take $G_1 = G_0$ and H_1 in the neighborhood of H_0 that form a feasible point of (37). Denote $\Delta H = H_1 - H_0$ and $e_\alpha = (b_\alpha - b_\alpha^0) - (\langle C_\alpha, G_0\rangle - \langle C_\alpha^0, G_0\rangle)$, then ΔH must satisfy

$$\langle B_\alpha, \Delta H\rangle = e_\alpha, |\alpha| \le 4$$

These set of equality constraints, which are equivalent to

$$\begin{aligned} \nu_2(z)' \Delta H \nu_2(z) &= \sum_{|\alpha|\le 4} e_\alpha z^\alpha \\ &= q(z) - q_0(z) - \nu_1(z)' G_0 \nu_1(z)(z'P'Pz - z'z), \end{aligned}$$

also implies that $\nu_2(e_i)' \Delta H \nu_2(e_i) = 0, i = 1, \ldots, r$. Therefore, $\{\nu_2(e_i) : i = 1, 2, \ldots, r\}$ spans a subspace of the null spaces of H_0, H_1 and ΔH. Since the null space of H_0 is exactly $\mathrm{span}(\{\nu_2(e_i) : i = 1, 2, \ldots, r\})$, and the minimal non-zero eigenvalue of H_0 is strictly greater than $1/3$ according to Theorem 4, it suffices to find a symmetric ΔH that satisfies

$$\langle B_\alpha, \Delta H\rangle = e_\alpha, |\alpha| \le 4$$

and $\|\Delta H\|$ is very small, much smaller than $\frac{1}{3}$.

In the following, we will complete the argument by showing that the solution $\Delta \hat{H}$ to

$$\begin{aligned} &\text{minimize } \|\Delta H\|_F \\ &\text{subject to } \langle B_\alpha, \Delta H\rangle = e_\alpha, |\alpha| \le 4. \end{aligned} \tag{38}$$

satisfies $\|\Delta H\|_F \le 0.0048$ under the conditions of $\|P'P - I\| \le 0.0016$, implying that $\Delta \bar{H} = \frac{1}{2}(\Delta \hat{H} + \Delta \hat{H}')$ is the desired ΔH.

We first estimate $\|e\|_\infty$. Note

$$\begin{aligned} q(z) - q_0(z) &= \sum_{k=1}^{r} [3(z'P'Pe_k) - 2z_k] z_k^2 - \sum_{k=1}^{r} z_k^3 \\ &= 3 \sum_{k=1}^{r} [(z'P'Pe_k) - z_k] z_k^2 \end{aligned}$$

which involves only third order monomials in sets $\{z_k^3 : k = 1, \ldots, r\}$, $\{z_k^2 z_j : k = 1, \ldots, r; j = r+1, \ldots, n\}$, and $\{z_k^2 z_j : j \ne k = 1, \ldots, r\}$. The coefficient for z_k^3 is $3(1 - e_k' P'Pe_k) = 0$, and the coefficient for $z_k^2 z_j$ is $-3e_j' P'Pe_k$. When $k = 1, \ldots, r; j = r+1, \ldots, n$, we have $-3e_j' P'Pe_k = 0$ due to the construction of P; when $j \ne k = 1, \ldots, r$, the quantity $-3e_j' P'Pe_k$ is non-zero. Therefore, we get

$$\|b - b^0\|_\infty \le 3 \max_{1 \le j \ne k \le r} |e_j' P'Pe_k| \le 3\epsilon.$$

We next bound

$$\begin{aligned} |\langle C_\alpha, G_0\rangle - \langle C_\alpha^0, G_0\rangle| &= |\langle C_\alpha - C_\alpha^0, G_0\rangle| \\ &= \frac{3}{2} |\operatorname{trace}(C_\alpha - C_\alpha)| \end{aligned}$$

To control $\operatorname{trace}(C_\alpha - C_\alpha^0)$, we write

$$\sum_{|\alpha|\le 4} (C_\alpha - C_\alpha^0) z^\alpha = \nu_1(z)\nu_1(z)'[z'(P'P - I)z]$$

Taking trace on both sides gives

$$\begin{aligned} &\sum_{|\alpha|\le 4} \operatorname{trace}(C_\alpha - C_\alpha^0) z^\alpha \\ &= \operatorname{trace}(\nu_1(z)\nu_1(z)')[z'(P'P - I)z] \\ &= \left(1 + \sum_{i=1}^{r} z_i^2\right)[z'(P'P - I)z] \\ &= \left(1 + \sum_{i=1}^{r} z_i^2\right) \sum_{1 \le j \ne k \le r} (P'P - I)_{jk} z_j z_k \end{aligned}$$

Since the diagonal of $P'P - I$ constitutes of zeros, the only monomials that have non-zero coefficients are in the sets $\{z_i^2 z_j z_k : 1 \le i \le r, 1 \le j \ne k \le r\}$, and $\{z_j z_k : 1 \le j \ne k \le r\}$. To compute the coefficients for $z_i^2 z_j z_k$, we consider two separate cases. When $j = i$, the coefficient for the term $z_i^3 z_k$ is $(P'P - I)_{ik} + (P'P - I)_{ki}$. When $j \ne i$ and $k \ne i$, the coefficient for the term $z_i^2 z_j z_k$ is $(P'P - I)_{jk} + (P'P - I)_{kj}$. In both cases, we can bound the absolute value of the coefficient by

$$\max_{j\neq k} |(P'P - I)_{jk} + (P'P - I)_{kj}| \leq 2\epsilon.$$

A similar argument shows that the coefficients for $z_j z_k$ with $1 \leq j \neq k \leq r$ are also bounded by 2ϵ. Hence, we get

$$\max_{|\alpha|\leq 4} |\operatorname{trace}(C_\alpha - C_\alpha^0)| \leq 2\epsilon.$$

Since the components of $b_\alpha - b_\alpha^0$ and $\langle C_\alpha - C_\alpha^0, G_0\rangle$ attain non-zero at different αs, we conclude that

$$\|e\|_\infty \leq 3\epsilon.$$

Denote by $S \in \mathbb{R}^{|\mathbb{N}_4^n|\times|\mathbb{N}_2^n|^2}$ the matrix whose αth row is $\operatorname{vec}(B_\alpha)^T$ for $|\alpha| \leq 4$. The solution to (38) is given by $\operatorname{vec}(\Delta H) = S^\dagger e$ where we used $\dagger$ to represent pseudo-inverse.

We want to control

$$\|S^\dagger\|_{\infty,2} = \max_\alpha \|[S^\dagger]_\alpha\|_2$$

where $[S^\dagger]_\alpha$ is the αth row of $S^\dagger$. Note S has orthogonal rows, and each $\operatorname{vec}(B_\alpha)$ is composed of zeros and ones, and the ones indicate where the monomial z^α locates in $\nu_2(z)\nu_2(z)'$. As a consequence, the matrix SS' is diagonal with the diagonal element d_α counts the number of appearances of z^α in $\nu_2(z)\nu_2(z)'$, which is always greater than or equal to 1. Therefore, we get

$$\begin{aligned}\|S^\dagger\|_{\infty,2} &= \|S'(SS')^{-1}\|_{\infty,2}\\ &\leq \max_\beta \|[S']_\beta \operatorname{diag}(d^{-1})\|_2\\ &\leq \max_\beta \|S^\beta\|_2\end{aligned}$$

where S^β represents that βth column of S. The index β indexes the rows and columns of $\nu_2(z)\nu_2(z)'$. Each column of S consists of zeros and a single one, with the latter representing which z^α is at the entry of $\nu_2(z)\nu_2(z)'$ specified by the column index β. Therefore, we obtain

$$\|S^\dagger\|_{\infty,2} \leq \max_\beta \|S^\beta\|_2 \leq 1$$

We conclude that

$$\begin{aligned}\|\Delta\bar{H}\|_F &\leq \|\Delta\hat{H}\|_F\\ &= \|S^\dagger e\|_2 \leq \|S^\dagger\|_{\infty,2}\|e\|_\infty\\ &\leq 3\epsilon\\ &\leq 0.0048\end{aligned}$$

for $\epsilon \leq 0.0016$. Therefore, the minimal non-zero eigenvalue of the Gram matrix $H_1 = H_0 + \Delta\bar{H}$ is lower bounded by $1/3 - 0.0048 > 0$.

So far we have showed that $q(z)$ is an SOS modulo the ellipsoid $\{z : z'P'Pz = 1\}$. To prove Theorem 1, we need to map z back to x, and make sure that after the mapping, the new Gram matrix still have rank $|\mathbb{N}_2^n| - r$. It suffices to show that the change of basis transformation on $\mathbb{R}^n$ that maps x to z induces a well-conditioned transformation between $\nu_2(x)$ and $\nu_2(z)$. This is given in Lemma 1 developed in the next section. Therefore, we have completed the proof of Theorem 1.

7.6. Change of Basis Formular

Consider two n-dimensional variables x and z linked by a change of basis transformation $x = Pz$ or $z = P^{-1}x$. We aim at finding the matrix J that expresses $\nu_2(z)$ in terms of $\nu_2(x)$, *i.e.*,

$$\nu_2(z) = \nu_2(P^{-1}x) = J\nu_2(x).$$

The transform J is well defined since a polynomial of degree k in z is always transformed into a polynomial of degree k in x under $z = P^{-1}x$. It's easy to see J has the form:

$$J = \begin{bmatrix} 1 & & \\ & P^{-1} & \\ & & J_2 \end{bmatrix}$$

where J_2 expresses all quadratic monomials of z in terms of quadratic monomials of x. To find J_2, we rewrite the relationship $zz' = P^{-1}xx'P^{-1'}$ as

$$\operatorname{vec}(zz') = P^{-1} \otimes_K P^{-1} \operatorname{vec}(xx')$$

where the subscript in the Kronecker product notation $\otimes_K$ is used to distinguish it from the tensor product notation $\otimes$, and $\operatorname{vec}(\cdot)$ vectorizes a matrix column-wise. Denote by $\bar{\nu}_2(x)$ all unique quadratic monomials in x, and write $\bar{\nu}_2(x) = \Pi \operatorname{vec}(xx')$, where Π is the matrix that picks and averages the duplicated quadratic monomials of x in $\operatorname{vec}(xx')$. One can verify that $\operatorname{vec}(xx') = \Pi^\dagger \bar{\nu}_2(x)$, and the smallest and largest singular values of Π are $\frac{1}{\sqrt{2}}$ and 1 respectively. Consequently, we have

$$\bar{\nu}_2(z) = \Pi \operatorname{vec}(zz') = \Pi \left(P^{-1} \otimes_K P^{-1}\right) \Pi^\dagger \bar{\nu}_2(x),$$

or equivalently $J_2 = \Pi P^{-1} \otimes_K P^{-1} \Pi^\dagger$. So if $\|PP' - I\| \leq \epsilon$, the singular values of J_2 are lower bounded and upper bounded by $\frac{1}{\sqrt{2}}\frac{1}{1+\epsilon}$ and $\frac{\sqrt{2}}{1-\epsilon}$ respectively. The same holds for J. We summarize these results in the following lemma.

Lemma 1. *The change of basis transformation $x = Pz$ induces a linear transformation between $\nu_2(z)$ and $\nu_2(x)$*

$$\nu_2(z) = J\nu_2(x) = \begin{bmatrix} 1 & & \\ & P^{-1} & \\ & & \Pi\left(P^{-1} \otimes_K P^{-1}\right)\Pi^\dagger \end{bmatrix} \nu_2(x)$$

such that the singular values of J fall into the interval $[\frac{1}{\sqrt{2}}\frac{1}{1+\epsilon}, \frac{\sqrt{2}}{1-\epsilon}]$.