Bounds for the estimator are in the paper, but here's a reminder of what this estimator is doing.

Suppose we compute $V = \frac{1}{\sqrt{k}} XR$, and look at any two rows i, j of V. Each tuple (v_{ik}, v_{jk}) are seen as bivariate normals, i.e.

$$\begin{pmatrix} v_{ik} \\ v_{jk} \end{pmatrix} \sim N(\mu, \Sigma)$$

with:

$$\mu = (0,0)$$
$$\Sigma = \frac{1}{K} \begin{pmatrix} m_i & a \\ a & m_j \end{pmatrix}$$

with a denoting the true inner product between $\mathbf{x}_i, \mathbf{x}_j$, and m_i, m_j denoting the norms of $\|\mathbf{x}_i\|_2^2$, $\|\mathbf{x}_j\|_2^2$ respectively.

For K columns of R, we see K such observations, and thus the likelihood is simply proportional to:

$$L((v_{ik}, v_{jk})) \propto |\Sigma|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \sum_{k=1}^{K} \left((v_{ik} \quad v_{jk}) \Sigma^{-1} \begin{pmatrix} v_{ik} \\ v_{jk} \end{pmatrix} \right) \right\}$$

The trick is to express the likelihood function in terms of a, the inner product, and find $\hat{a}$ which maximizes this likelihood.

The loglikelihood (derivation in the paper) is given by:

$$l(a) = -\frac{K}{2} \log(m_i m_j - a^2) - \frac{K}{2} \frac{1}{m_i m_j - a^2} \sum_{k=1}^{K} (v_{ik}^2 m_j - 2 v_{ik} v_{jk} a + v_{jk}^2 m_i)$$

Thus, getting the MLE of a is equivalent to equating $l'(a) = 0$, and finding $\hat{a}$ which gives $l'(\hat{a}) = 0$. Thus, need root finding code to do this. $\hat{a}$ solution to:

$$a^3 - a^2(\mathbf{v}_i^T \mathbf{v}_j) + a(-m_i m_j + m_i \|\mathbf{v}_j\|_2^2 + m_j \|\mathbf{v}_i\|_2^2) - m_i m_j \mathbf{v}_i^T \mathbf{v}_j = 0$$