# Appendix

## Table of Contents



## A Discussion and outlook

**Intuitions on the separation rate** Let us provide some explanations that should help to gain intuition on the conditions on $\bar\kappa_{\text{in-in}}$ and $\bar\kappa_{\text{in-out}}$ obtained in our main theorems. More precisely, we will explain in this paragraph where the right hand side of (6) comes from. Consider the simpler problem in which we wish to test the hypothesis $H_0 : \boldsymbol{\mu} = 0$ against $H_1 : \boldsymbol{\mu} \neq 0$ based on the observation $\boldsymbol{Y}$ drawn from the Gaussian distribution $\mathcal{N}_d(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_d)$. This problem has a tight link with the considered problem of matching, since one can think of $\boldsymbol{Y}$ as the difference $\boldsymbol{X}_i - \boldsymbol{X}^{\#}_j$. We are interested in checking whether the pair $(i, j)$ is such that $j = \pi^*(i)$, that is whether $H_0$ is true.

Using the standard bounds on the tails of the chi-squared distribution (Lemma 1), one can check that under $H_0$, the random vector $\boldsymbol{Y}$ lies with probability $\geq 1 - \alpha$ in the ring $\mathfrak{R}_0 = B(0, \sigma\sqrt{d + r_2}) \setminus B(0, \sigma\sqrt{d - r_1})$ where

$$r_1 = 2\sqrt{d \log(1/\alpha)} \qquad \text{and} \qquad r_2 = 2\sqrt{d \log(1/\alpha)} + 2\log(1/\alpha).$$

Similarly, considering the approximation $\|\boldsymbol{Y}\|_2^2 \approx \|\boldsymbol{\mu}\|_2^2 + \sigma^2\|\boldsymbol{\xi}\|_2^2$ where $\boldsymbol{\xi}$ is a standard Gaussian vector, we can check that under $H_1$, the random vector $\boldsymbol{Y}$ lies with probability $\geq 1 - \alpha$ in the ring $\mathfrak{R}_1 = B(0, \sigma\sqrt{\|\boldsymbol{\mu}/\sigma\|_2^2 + d + r_2}) \setminus B(0, \sigma\sqrt{\|\boldsymbol{\mu}/\sigma\|_2^2 + d - r_1})$.

If the two rings $\mathfrak{R}_0$ and $\mathfrak{R}_1$ are disjoint, it is possible to decide between $H_0$ and $H_1$ by checking whether $\boldsymbol{Y}$ belongs to $\mathfrak{R}_0$ or not. This condition of disjointness is equivalent to

$$\|\boldsymbol{\mu}/\sigma\|_2^2 + d - r_1 > d + r_2.$$

This leads to

$$\begin{aligned}\|\boldsymbol{\mu}/\sigma\|_2 > \sqrt{r_1 + r_2} &= \big(4\sqrt{d\log(1/\alpha)} + 2\log(1/\alpha)\big)^{1/2}\\ &\asymp \big(d\log(1/\alpha)\big)^{1/4} \vee \log^{1/2}(1/\alpha).\end{aligned}$$

The right hand side of the last display is of the same order as the right hand side of the (6), for small values of $nm$. The fact that for large values of $nm$ there is a logarithmic deterioration, due to the fact that we have to test a large number of hypotheses $H_{0,i,j} : \theta_{\pi^*(i)} = \theta^{\#}_j$, $(i, j) \in [n] \times [m]$, is quite common in probability and statistics.

**Other noise distributions** The results of this paper can be extended to sub-Gaussian distributions without any change in the rates. The extension to sub-exponential distributions seems also possible to do using the methodology employed in this paper, but will most likely lead to higher-order polylogarithmic terms.

Finally, considering heavy tailed distributions such as the multivariate Student distribution might have stronger impact on the rate. Studying this impact is out of scope of the present work.

**Outlier detection** The results presented in previous sections provide conditions under which the objective mapping is identified with high probability. This automatically implies that the outliers are correctly identified. However, the task of outlier detection is arguably simpler than that of estimation of $\pi^*$. Therefore, one may wonder whether this task can be accomplished under weaker assumptions than those required in the theorems stated in this paper. Somewhat surprisingly, it turns out that this is not the case unless we require the outliers to be very far away from the inliers.

Indeed, on the one hand, if the normalized distance between the outliers and the inliers is not larger than $O(d^{1/2})$, it follows from the counter-example constructed in the proof of Theorem 3 that it is impossible to identify the outliers using a distance based $M$-estimator. Extending the arguments presented in Appendix C below, one can check that this impossibility holds for every estimator of the set of outliers.

On the other hand, suitably adapting the arguments of the proof of Theorem 4, one can prove that if the inlier-outlier distance is larger than a threshold of order $\sqrt{d}\exp(cn)$ for some $c > 0$, the LSL recovers the true set of outliers.

**Estimation of $\pi^*$ instead of detection** An interesting yet challenging problem is that of assessing the minimax risk of estimation of $\pi^*$ when the error is measured, for instance, by means of the Hamming loss $\ell_{\text{Hamming}}(\hat{\pi};\pi^*) = \#\{i \in [n] : \hat{\pi}(i) \neq \pi^*(i)\}$. It is relevant to study this problem in a setting where consistent detection of $\pi^*$ (*i.e.*, Hamming loss equal to zero) is impossible, that is when the separation conditions are violated but some weaker assumptions are satisfied. On a related note, one may look for conditions on the normalized separation distances which ensure the existence of an estimator $\hat{\pi}$ such that $\mathbf{P}(\ell_{\text{Hamming}}(\hat{\pi};\pi^*) \leq \tau n) \geq 1-\alpha$. This means that with probability $\geq 1-\alpha$ the fraction of mismatched vectors of the estimated map $\hat{\pi}$ is less than $\tau$, for $\tau \in (0,1)$. Note that these problems are not studied even in the simpler outlier-free framework.

# B Postponed proofs

In this appendix we have collected the proofs of the theorems presented in the main text of the paper, as well as some technical definitions used in the proofs. First, denote

$$\sigma_{i,j}^2 = \sigma_i^2 + \sigma_j^{\#2} \qquad \text{and} \qquad \kappa_{i,j} = \frac{\|\theta_i - \theta_j^\#\|}{\sigma_{i,j}} \tag{11}$$

for any pair of indices $(i,j)$ with $i \in [n]$ and $j \in [m]$. We will also use the notation

$$\bar{\kappa} = \min(\bar{\kappa}_{\text{in-in}}, \bar{\kappa}_{\text{in-out}}). \tag{12}$$

Second, we define the random variables $\zeta_1$ and $\zeta_2$ as follows

$$\zeta_1 = \max_{i\neq j} \frac{\big|(\theta_i - \theta_j^\#)^\top(\sigma_i\xi_i - \sigma_j^\#\xi_j^\#)\big|}{\|\theta_i - \theta_j^\#\|\sigma_{i,j}}, \quad \zeta_2 = d^{-1/2}\max_{i,j}\left|\frac{\big\|\sigma_i\xi_i - \sigma_j^\#\xi_j^\#\big\|^2}{\sigma_{i,j}^2} - d\right|.$$

It can be easily noticed that $\zeta_1 = \max_{i\neq j}|\zeta_{i,j}|$, where $\zeta_{i,j}$ are standard Gaussian random variables. As for $\zeta_2$, it can be seen that $\zeta_2 = d^{-1/2}\max_{i,j}|\eta_{i,j}|$, where $\eta_{i,j}$ are centered $\chi^2$ random variables with $d$ degrees of freedom, *i.e.* $\eta_{i,j} \overset{\mathscr{D}}{=} \chi_d^2 - d$.

In addition, one can infer from (1) that for every $i \in [n]$ and every $j \in [m]$, we have

$$\begin{aligned}
\|X_i - X_j^\#\|^2 &\leq \|\theta_i - \theta_j^\#\|^2 + \sigma_{i,j}^2(d + \sqrt{d}\,\zeta_2) + 2\zeta_1\|\theta_i - \theta_j^\#\|\sigma_{i,j} \\
&= \sigma_{i,j}^2\big(\kappa_{i,j}^2 + d + \sqrt{d}\,\zeta_2 + 2\zeta_1\kappa_{i,j}\big), \qquad (13)\\
\|X_i - X_j^\#\|^2 &\geq \|\theta_i - \theta_j^\#\|^2 + \sigma_{i,j}^2(d - \sqrt{d}\,\zeta_2) - 2\zeta_1\|\theta_i - \theta_j^\#\|\sigma_{i,j} \\
&= \sigma_{i,j}^2\big(\kappa_{i,j}^2 + d - \sqrt{d}\,\zeta_2 - 2\zeta_1\kappa_{i,j}\big). \qquad (14)
\end{aligned}$$

The concentration of the centered and normalized $\chi^2$ random variable, such as $\zeta_2$, is described in the following lemma.

**Lemma 1 (Laurent and Massart (2000), Eq. (4.3) and (4.4))** *If $Y$ is drawn from the chi-squared distribution $\chi^2(D)$, where $D \in \mathbb{N}^*$, then, for every $x > 0$,*

$$\begin{cases} \mathbf{P}\big(Y - D \le -2\sqrt{Dx}\big) \le e^{-x}, \\ \mathbf{P}\big(Y - D \ge 2\sqrt{Dx} + 2x\big) \le e^{-x}. \end{cases}$$

*As a consequence, for every $y > 0$, $\mathbf{P}\big(D^{-1/2}|Y - D| \ge y\big) \le 2\exp\big\{ -\frac{1}{8}y(y \wedge \sqrt{D})\big\}$. Or, equivalently, for any $\alpha \in (0,1)$, we have*

$$\mathbf{P}\bigg(D^{-1/2}|Y - D| \le 2\sqrt{\log(2/\alpha)} + \frac{2\log(2/\alpha)}{\sqrt{D}}\bigg) \ge 1 - \alpha.$$

### B.1 Proof of Theorem 1

We prove the upper bound for $\bar\kappa$ in the presence of outliers. Without loss of generality we can assume that $\pi^*(i) = i, \quad \forall i \in [n]$. We wish to bound the probability of the event $\Omega = \{\hat\pi \neq \pi^*\}$, where $\hat\pi = \bar\pi^{\mathrm{LSNS}}$. It is evident that

$$\Omega \subset \bigcup_{\pi \neq \pi^*} \Omega_\pi, \tag{15}$$

where the union is taken over all possible injective mappings $\pi : [n] \to [m]$ and

$$\Omega_\pi = \bigg\{ \sum_{i=1}^n \frac{\|X_i - X_i^\#\|^2}{2\sigma_i^2} \ge \sum_{i=1}^n \frac{\|X_i - X_{\pi(i)}^\#\|^2}{\sigma_i^2 + (\sigma_{\pi(i)}^\#)^2} \bigg\}.$$

One easily checks that the following inclusion holds:

$$\Omega_\pi \subset \bigcup_{i=1}^n \bigcup_{j \in [m]\setminus\{i\}} \bigg\{ \frac{\|X_i - X_i^\#\|^2}{2\sigma_i^2} \ge \frac{\|X_i - X_j^\#\|^2}{\sigma_i^2 + (\sigma_j^\#)^2} \bigg\}. \tag{16}$$

Since $\pi^*(i) = i$ for every $i \in [n]$, $\kappa_{i,i} = 0$ (see the definition in (11)) and, in view of (13),

$$\|X_i - X_i^\#\|^2 \le 2\sigma_i^2(d + \sqrt{d}\,\zeta_2). \tag{17}$$

Similarly, for every $j \in [m]$ and $j \neq i$, in view of (14),

$$\|X_i - X_j^\#\|^2 \ge \sigma_{i,j}^2(\kappa_{i,j}^2 + d - \sqrt{d}\,\zeta_2 - 2\kappa_{i,j}\zeta_1).$$

Recall that $\bar\kappa$ defined in (12), is the smallest normalized distance $\kappa_{i,j}$. Therefore, on the event $\Omega_1 = \{\bar\kappa \ge \zeta_1\}$, the previous display implies that

$$\frac{\|X_i - X_j^\#\|^2}{\sigma_{i,j}^2} \ge \bar\kappa^2 - 2\bar\kappa\zeta_1 + d - \sqrt{d}\,\zeta_2. \tag{18}$$

Hence, combining obtained bounds (17) and (18) we get that

$$\begin{aligned} \bigg\{ \frac{\|X_i - X_i^\#\|^2}{2\sigma_i^2} \ge \frac{\|X_i - X_j^\#\|^2}{\sigma_i^2 + (\sigma_j^\#)^2} \bigg\} \cap \Omega_1 &\subset \Big\{ d + \sqrt{d}\,\zeta_2 \ge \bar\kappa^2 - 2\bar\kappa\zeta_1 + d - \sqrt{d}\,\zeta_2 \Big\} \\ &= \Big\{ 2\sqrt{d}\,\zeta_2 + 2\bar\kappa\,\zeta_1 \ge \bar\kappa^2 \Big\}. \end{aligned} \tag{19}$$

Note that the event on the right hand side of the last display is independent of the pair $(i, j)$. This implies that

$$\begin{aligned} \Omega \cap \Omega_1 &\overset{\text{by (15)}}{\subset} \bigg( \bigcup_{\pi \neq \pi^*} \Omega_\pi \bigg) \cap \Omega_1 \\ &\overset{\text{by (16)}}{\subset} \bigg( \bigcup_{i=1}^n \bigcup_{j \in [m]\setminus\{i\}} \bigg\{ \frac{\|X_i - X_i^\#\|^2}{2\sigma_i^2} \ge \frac{\|X_i - X_j^\#\|^2}{\sigma_i^2 + (\sigma_j^\#)^2} \bigg\} \bigg) \cap \Omega_1 \\ &\subset \bigcup_{i=1}^n \bigcup_{j \in [m]\setminus\{i\}} \bigg( \bigg\{ \frac{\|X_i - X_i^\#\|^2}{2\sigma_i^2} \ge \frac{\|X_i - X_j^\#\|^2}{\sigma_i^2 + (\sigma_j^\#)^2} \bigg\} \cap \Omega_1 \bigg) \\ &\overset{\text{by (19)}}{\subset} \Big\{ 2\sqrt{d}\,\zeta_2 + 2\bar\kappa\,\zeta_1 \ge \bar\kappa^2 \Big\}. \end{aligned} \tag{20}$$

Using (20) we can show that

$$\begin{aligned}\mathbf{P}(\Omega) &\le \mathbf{P}(\Omega_1^{\complement}) + \mathbf{P}\big(\Omega\cap\Omega_1\big)\\ &\le \mathbf{P}\big(\zeta_1 \ge \bar\kappa\big) + \mathbf{P}\big(2\sqrt{d}\zeta_2 + 2\bar\kappa\zeta_1 \ge \bar\kappa^2\big)\\ &\le \mathbf{P}\big(\zeta_1 \ge \bar\kappa\big) + \mathbf{P}\big(\zeta_1 \ge \tfrac14\bar\kappa\big) + \mathbf{P}\Big(2\sqrt{d}\zeta_2 + 2\bar\kappa\zeta_1 \ge \bar\kappa^2; \zeta_1 < \tfrac14\bar\kappa\Big)\\ &\le 2\mathbf{P}\big(\zeta_1 \ge \tfrac14\,\bar\kappa\big) + \mathbf{P}\Big(\zeta_2 \ge \frac{\bar\kappa^2}{4\sqrt{d}}\Big). \end{aligned} \tag{21}$$

For suitably chosen standard Gaussian random variables $\zeta_{i,j}$ it holds that $\zeta_1 = \max_{i\neq j}|\zeta_{i,j}|$. Therefore, using the tail bound for the standard Gaussian distribution and the union bound, we get

$$\mathbf{P}\Big(\zeta_1 \ge \tfrac14\bar\kappa\Big) \le \sum_{i\neq j}\mathbf{P}\Big(|\zeta_{i,j}| \ge \tfrac14\bar\kappa\Big) \le 2nm\,e^{-\bar\kappa^2/32}.$$

To complete the proof, it remains to upper bound the second term in the right hand side of (21), *i.e.*, to evaluate the tail of the random variable $\zeta_2$. To this end, we use the concentration result stated in Lemma 1 with $y = \frac{\bar\kappa^2}{4\sqrt{d}}$, combined with the union bound and simple algebra. This yields

$$\begin{aligned}\mathbf{P}\Big(\zeta_2 \ge \frac{\bar\kappa^2}{4\sqrt{d}}\Big) &\le 2nm\exp\Big\{-\frac18\cdot\frac{\bar\kappa^2}{4\sqrt{d}}\Big(\frac{\bar\kappa^2}{4\sqrt{d}}\wedge\sqrt{d}\Big)\Big\}\\ &= 2nm\exp\Big\{-\frac{(\bar\kappa/16)^2}{d}(2\bar\kappa^2\wedge 8d)\Big\},\end{aligned} \tag{22}$$

where the $nm$ factor in front of the exponent comes from the union bound for all $nm$ pairs $(i,j)$ from the definition of $\zeta_2$, while the exponent is a direct application of Lemma 1. Finally, using inequalities (21)-(22), we get that whenever

$$\bar\kappa \ge 4\Big(\sqrt{2\log(8nm/\alpha)} \vee \big(d\log(4nm/\alpha)\big)^{1/4}\Big), \tag{23}$$

the probability of incorrect matching is at most $\alpha$. Thus, we have formally showed that if (23) holds then $\mathbf{P}(\hat\pi \neq \pi^*) = \mathbf{P}(\Omega) \le \alpha$, as desired.

### B.2 Proof of Theorem 2

We prove the upper bound for $\bar\kappa = \min(\bar\kappa_{\text{in-in}}, \bar\kappa_{\text{in-out}})$ in the presence of outliers and in the case of unknown noise variance. We wish to bound the probability of the event $\Omega = \{\hat\pi \neq \pi^*\}$, where $\hat\pi = \hat\pi^{\text{LSL}}$ and $\pi^*(i) = i$ for all $i \in [n]$. It is evident that

$$\Omega \in \bigcup_{\pi\neq\pi^*}\Omega_\pi, \tag{24}$$

where

$$\begin{aligned}\Omega_\pi &= \Big\{\sum_{i=1}^n \log\|X_i - X_i^{\#}\|^2 \ge \sum_{i=1}^n \log\|X_i - X_{\pi(i)}^{\#}\|^2\Big\}\\ &\subset \bigcup_{i=1}^n\bigcup_{j\in[m]\setminus\{i\}}\Big\{\log\|X_i - X_i^{\#}\|^2 \ge \log\|X_i - X_j^{\#}\|^2\Big\}\end{aligned} \tag{25}$$

Recall that $\bar\kappa = \min(\bar\kappa_{\text{in-in}}, \bar\kappa_{\text{in-out}})$. On the event $\Omega_1 = \{\bar\kappa \ge \zeta_1\}$, from (14), we get

$$\frac{\|X_i - X_j^{\#}\|^2}{\sigma_{i,j}^2} \ge \bar\kappa^2 - 2\zeta_1\bar\kappa + d - \sqrt{d}\,\zeta_2. \tag{26}$$

Note that the expression on the right hand side of the last display is independent of the pair $(i,j)$. This implies that

$$\begin{aligned}
\Omega\cap\Omega_1 &\subset \Big(\bigcup_{\pi\neq\pi^*}\Omega_\pi\Big)\cap\Omega_1 && \text{[by (24)]}\\
&\subset \Big(\bigcup_{i=1}^n\bigcup_{j\in[m]\setminus\{i\}}\Big\{\log\|X_i-X_i^\#\|^2\ge\log\|X_i-X_j^\#\|^2\Big\}\Big)\cap\Omega_1 && \text{[by (25)]}\\
&\subset \bigcup_{i=1}^n\bigcup_{j\in[m]\setminus\{i\}}\Big(\Big\{\|X_i-X_i^\#\|^2\ge\|X_i-X_j^\#\|^2\Big\}\cap\Omega_1\Big)\\
&\subset\Big\{2\sigma_i^2(d+\sqrt{d}\zeta_2)\ge\sigma_{i,j}^2(\bar\kappa^2-2\zeta_1\bar\kappa+d-\sqrt{d}\,\zeta_2)\Big\} && \text{[by (13),(26)]}\\
&\subset\Big\{2(d+\sqrt{d}\zeta_2)\ge\bar\kappa^2-2\zeta_1\bar\kappa+d-\sqrt{d}\,\zeta_2\Big\}, && \text{[since } \sigma_i\le\sigma_{i,j}]\\
&\subset\Big\{3\sqrt{d}\,\zeta_2+2\zeta_1\bar\kappa\ge\bar\kappa^2-d\Big\}. && (27)
\end{aligned}$$

We can bound the probability of incorrect matching $\mathbf{P}(\Omega)$ using the relationship obtained in (27)

$$\begin{aligned}
\mathbf{P}(\Omega)&\le\mathbf{P}(\Omega_1^\complement)+\mathbf{P}(\Omega\cap\Omega_1)\\
&\le\mathbf{P}(\zeta_1\ge\bar\kappa)+\mathbf{P}\big(3\sqrt{d}\,\zeta_2+2\zeta_1\bar\kappa\ge\bar\kappa^2-d\big).
\end{aligned}$$

From the last inequality, we infer that

$$\begin{aligned}
\mathbf{P}(\Omega)&\le\mathbf{P}(\zeta_1\ge\bar\kappa)+\mathbf{P}\Big(\zeta_1\ge\tfrac14\bar\kappa\Big)+\mathbf{P}\Big(3\sqrt{d}\,\zeta_2+2\zeta_1\bar\kappa\ge\bar\kappa^2-d\,;\,\zeta_1<\tfrac14\bar\kappa\Big)\\
&\le 2\mathbf{P}\Big(\zeta_1\ge\tfrac14\bar\kappa\Big)+\mathbf{P}\Big(3\sqrt{d}\,\zeta_2\ge\tfrac12\bar\kappa^2-d\Big)\\
&\le 2\mathbf{P}\Big(\zeta_1\ge\tfrac14\bar\kappa\Big)+\mathbf{P}\Big(\zeta_2\ge\frac{\bar\kappa^2-2d}{6\sqrt{d}}\Big). \qquad (28)
\end{aligned}$$

As mentioned in the beginning of the section, for suitably chosen standard Gaussian random variables $\zeta_{i,j}$ it holds that $\zeta_1=\max_{i\neq j}|\zeta_{i,j}|$. Therefore, using the tail bound for the standard Gaussian distribution and the union bound, we get

$$\mathbf{P}\Big(\zeta_1\ge\tfrac14\bar\kappa\Big)\le\sum_{i\neq j}\mathbf{P}\Big(|\zeta_{i,j}|\ge\tfrac14\bar\kappa\Big)\le 2nm\,e^{-\bar\kappa^2/32}\le\alpha/4. \qquad (29)$$

To complete the proof, it remains to upper bound the second term in the right hand side of (28), *i.e.*, to evaluate the tail of the random variable $\zeta_2$. Using Lemma 1 with $y=(\bar\kappa^2-2d)/(6\sqrt{d})$—which is positive under the conditions of the theorem—combined with the union bound, we arrive at

$$\begin{aligned}
\mathbf{P}(\zeta_2\ge y)&\le 2nm\exp\Big\{-\frac18 y(y\wedge\sqrt{d})\Big\}\\
&=2nm\Big(\exp\Big\{-\frac18 y^2\Big\}\vee\exp\Big\{-\frac18 y\sqrt{d}\Big\}\Big).
\end{aligned}$$

One easily checks that the last expression is smaller than $\alpha/2$ if and only if

$$y^2\ge 8\log(4nm/\alpha)\qquad\text{and}\qquad y\sqrt{d}\ge 8\log(4nm/\alpha)$$

which is equivalent to

$$y\ge\big(2\sqrt{2\log(4nm/\alpha)}\big)\vee\big((8/\sqrt{d})\log(4nm/\alpha)\big).$$

Replacing $y=(\bar\kappa^2-2d)/(6\sqrt{d})$, the last inequality becomes

$$\bar\kappa^2\ge 2d+\big(12\sqrt{2d\log(4nm/\alpha)}\big)\vee\big(48\log(4nm/\alpha)\big).$$

Combining the inequality from the last display with the bound derived from (29) we get that all these bounds are satisfied whenever

$$\bar\kappa\ge\sqrt{2d}+4\Big\{\Big(2d\log\frac{4nm}{\alpha}\Big)^{1/4}\vee\Big(3\log\frac{8nm}{\alpha}\Big)^{1/2}\Big\}.$$

Therefore, under this condition on $\bar\kappa$, the probability of the incorrect matching is at most $\alpha$, *i.e.* $\mathbf{P}(\hat\pi\neq\pi^*)=\mathbf{P}(\Omega)\le\alpha$.

### B.3 Proof of Theorem 3

First we fix $m = n+1$ and $\pi^*(i) = i$ for all $i \in [n]$, where $\pi^*$ is the correct matching. Let $\sigma_1^\# = 1$ and $\sigma_{i+1}^\# = \alpha^i$ for all $i \in [n]$, where $\alpha \ll 1$. Then let's take $\pi(i) = i+1$ for all $i \in [n]$. Let $L(\pi)$ be the vector of distances $\|X_i - X^\#_{\pi(i)}\|$ for a matching scheme $\pi$

$$L(\pi) = \begin{bmatrix} \|X_1 - X^\#_{\pi(1)}\| \\ \|X_2 - X^\#_{\pi(2)}\| \\ \dots \\ \|X_n - X^\#_{\pi(n)}\| \end{bmatrix}.$$

The next lemma shows that the event $L(\bar\pi) < L(\pi^*)$ (coordinate-wise) occurs with probability at least $1/4$.

**Lemma 2** *Let $n \geq 4$, $d \geq 422\log(4n)$ and $\theta_1^\# = (1; 0; \dots; 0)^\top$. Assume that $\pi^*(i) = i$, $\sigma_i^\# = 2^{-(i-1)}$ and $\theta^\#_{i+1} = \theta^\#_i + 2^{-(i+1)}\sqrt{d}\,\theta^\#_1$ for all $i \in [n+1]$. Then $L(\pi^*) > L(\bar\pi)$ with probability greater than $1/4$, where $\bar\pi$ is the injection defined by $\bar\pi(i) = i+1$. Furthermore, for these values $(\boldsymbol{\theta}^\#, \boldsymbol{\sigma}^\#, \pi^*)$, we have $\kappa_{\text{in-in}} = \kappa_{\text{in-out}} = \sqrt{d/20}$.*

**Proof of Lemma 2** Let us denote

$$\bar\kappa_i \triangleq \frac{\|\theta^\#_{\pi(i)} - \theta_i\|}{\sqrt{\sigma_i^2 + \sigma^{\#2}_{\bar\pi(i)}}} = \sqrt{d/20}, \qquad \text{for all } i \in [n].$$

Recall that $\sigma^2_{i,j} = \sigma_i^2 + \sigma_j^{\#2}$ and write

$$L_i(\pi) = \|X_i - X^\#_{\pi(i)}\|^2 = \|\theta_i - \theta^\#_{\pi(i)} + \zeta_i \sigma_{i,\pi(i)}\|^2,$$

where $\zeta_i \sim \mathcal{N}(0, I_d)$. Notice that $L_i(\pi^*) = 2\sigma_i^2\|\zeta_i\|^2$ for all $i \in [n]$. Similarly, the expression from the last display for $\bar\pi$ reads as

$$L_i(\bar\pi) = \|\zeta_i \sigma_{i,\bar\pi(i)}\|^2\Big(1 + \frac{\bar\kappa_i^2}{\|\zeta_i\|^2}\Big) + 2\sigma_{i,\bar\pi(i)}\zeta_i^\top(\theta_i - \theta^\#_{\bar\pi(i)}).$$

Plugging in the values of $\boldsymbol{\sigma}^\#$ with $\alpha = 1/2$ and $\bar\pi(i) = i+1$ we arrive at

$$L_i(\pi^*) = 2^{3-2i}\|\zeta_i\|^2, \qquad L_i(\bar\pi) = \frac{5}{2^{2i}}\|\bar\zeta_i\|^2\Big(1 + \frac{\bar\kappa_i^2}{\|\bar\zeta_i\|^2}\Big) + \frac{\sqrt{5}}{2^{i-1}}\bar\zeta_i^\top(\theta_i - \theta^\#_{i+1}),$$

where in the second expression we write $\bar\zeta_i$ instead of $\zeta_i$ to indicate that these random variables are different, though both are standard normal $d$-dimensional vectors. We first replace the second term of $L_i(\bar\pi)$ with its upper bound that holds with probability of at least $1/4$. It is evident that the random variable $Z \triangleq 2\sigma_{i,\bar\pi(i)}\zeta_i^\top(\theta_i - \theta^\#_{\bar\pi(i)})$ is Gaussian with standard deviation $\sigma \triangleq 2\sigma_{i,\bar\pi(i)}\|\theta_i - \theta^\#_{\bar\pi(i)}\| = 2\sigma^2_{i,\bar\pi(i)}\bar\kappa_i$, therefore

$$\mathbf{P}(Z \geq \sigma\sqrt{2\log 4}) \leq \frac{1}{4}.$$

Hence, on the event $\Omega = \{Z \leq 2\sigma^2_{i,\bar\pi(i)}\bar\kappa_i\sqrt{2\log 4}\}$ the inequality $L_i(\pi^*) > L_i(\bar\pi)$ holds whenever

$$\begin{aligned} &\frac{8}{2^{2i}}\|\zeta_i\|^2 > \frac{5}{2^{2i}}\|\bar\zeta_i\|^2\Big(1 + \frac{\bar\kappa_i^2}{\|\bar\zeta_i\|^2}\Big) + \frac{5}{2^{2i}}\bar\kappa_i\sqrt{8\log 4}, \\ &\frac{8}{5}\|\zeta_i\|^2 - \|\bar\zeta_i\|^2 > \bar\kappa_i^2 + 2\bar\kappa_i\sqrt{2\log 4}. \end{aligned} \tag{30}$$

Notice that the left hand side of (30) is a weighted difference of two centered and normalized $\chi^2$ random variables with $d$ degrees of freedom. The concentration inequality for such difference is a direct consequence of Lemma 1. Namely, for $X, Y \sim \chi^2_d$ the concentration bound for $Z = \alpha X - \beta Y$ with arbitrary $\alpha, \beta \in \mathbb{R}$ reads as

$$\mathbf{P}(Z \geq (\alpha - \beta)d - 2\sqrt{dx}(\alpha + \beta) - 2\beta x) \geq 1 - 2e^{-x}.$$

It is easy to verify that given $n \geq 4, d \geq 422\log(4n)$ and $\bar{\kappa}_i \leq \sqrt{d/20}$, then

$$\bar{\kappa}_i^2 + 2\bar{\kappa}_i\sqrt{2\log 4} \leq \frac{3}{5}d - \frac{26}{5}\sqrt{d\log(4n)} - 2\log(4n),$$

where the right hand side is the quantile of $Z$ with $x = \log(4n)$. Combining the inequality from the last display with (30) we get that on the event $\Omega$ we have

$$\mathbf{P}(L_i(\pi^*) > L_i(\bar{\pi})) \geq 1 - \frac{1}{2n}.$$

Recall that $\mathbf{P}(\Omega) \geq 3/4$, then using the union bound for events $\Omega$ and $\{L_i(\pi^*) > L_i(\bar{\pi})\}$ all $i \in [n]$ we arrive at $\mathbf{P}(L(\pi^*) > L(\bar{\pi})) > 1/4$. This completes the proof of Lemma 2.

Therefore, using the result of Lemma 2 and applying any non-decreasing function $\rho(\cdot)$ to each of the coordinates of $L(\bar{\pi})$ and $L(\pi^*)$ yields

$$\sum_{i=1}^{n}\rho_i(\|X_i - X^{\#}_{\bar{\pi}(i)}\|) < \sum_{i=1}^{n}\rho_i(\|X_i - X^{\#}_{\pi^*(i)}\|)$$

with probability of at least $1/4$. This, in turn, implies that an optimizer will not choose $\pi^*$ on this event. Hence, $\mathbf{P}(\bar{\pi} \neq \pi^*) > 1/4$, concluding the proof of the theorem.

### B.4 Proof of Theorem 4

To ease notation, we write $\hat{\pi}$ instead of $\hat{\pi}^{\mathrm{LSL}}_{n,m}$, and, without loss of generality, we assume that $\pi^*(i) = i$ for $i \in [n]$. We wish to prove that on an event of probability $\geq 1-\alpha$, for every injective mapping $\pi : [n] \to [m]$, we have $\psi(\pi^*) \leq \psi(\pi)$, where

$$\psi(\pi) = \sum_{i=1}^{n}\log\|X_i - X^{\#}_{\pi(i)}\|^2.$$

Since the logarithm is an increasing function, this is equivalent to showing that

$$\prod_{i=1}^{n}\|X_i - X^{\#}_{\pi^*(i)}\|^2 < \prod_{i=1}^{n}\|X_i - X^{\#}_{\pi(i)}\|^2, \quad \text{for every } \pi \neq \pi^*,$$

which, in turn, is the same as

$$\prod_{i=1}^{n}\frac{\|X_i - X^{\#}_{\pi^*(i)}\|^2}{\|X_i - X^{\#}_{\pi(i)}\|^2} < 1, \quad \text{for every } \pi \neq \pi^*.$$

In view of (13) and (14), we have

$$\begin{aligned}\prod_{i=1}^{n}\frac{\|X_i - X^{\#}_{\pi^*(i)}\|^2}{\|X_i - X^{\#}_{\pi(i)}\|^2} &\leq \prod_{\substack{i\in[n]\\ \pi(i)\neq\pi^*(i)}} \frac{2\sigma_i^2(d+\sqrt{d}\,\zeta_2)}{\sigma_{i,\pi(i)}^2\big(\kappa_{i,\pi(i)}^2 + d - \sqrt{d}\,\zeta_2 - 2\zeta_1\kappa_{i,\pi(i)}\big)_+}\\ &\leq \prod_{\substack{i\in[n]\\ \pi(i)\neq\pi^*(i)}} \frac{4\sigma_i^2(d+\sqrt{d}\,\zeta_2)}{\sigma_{i,\pi(i)}^2\big(\kappa_{i,\pi(i)}^2 + 2d - 2\sqrt{d}\,\zeta_2\big)_+}, \qquad \text{if} \quad \zeta_1 \leq (1/4)\bar{\kappa}. \quad (31)\end{aligned}$$

Let us define the sets $I_1 = \big\{i \in [n] : \pi(i) \in \mathrm{Im}(\pi^*)\backslash\{\pi^*(i)\}\big\}$ and $I_2 = \{i \in [n] : \pi(i) \notin \mathrm{Im}(\pi^*)\}$. Clearly, using the inequality $\sigma_{i,j}^2 \geq 2\sigma_i\sigma_j^{\#}$, we get

$$\prod_{\substack{i\in[n]\\ \pi(i)\neq\pi^*(i)}}\frac{2\sigma_i^2}{\sigma_{i,\pi(i)}^2} \leq \prod_{\substack{i\in[n]\\ \pi(i)\neq\pi^*(i)}}\frac{\sigma_i^2}{\sigma_i\sigma^{\#}_{\pi(i)}} = \frac{\prod_{i\in I_1\cup I_2}\sigma_i}{\prod_{i\in I_1}\sigma^{\#}_{\pi(i)}\prod_{i\in I_2}\sigma^{\#}_{\pi(i)}}. \quad (32)$$

For every $i \in I_1$, there is $j \in [n]$ such that $\pi(i) = \pi^*(j)$; this $j$ is given by $j = (\pi^*)^{-1}(i)$. For such a pair $(i, j)$, in view of (2), we have $\sigma^{\#}_{\pi(i)} = \sigma^{\#}_{\pi^*(j)} = \sigma_j$. Note that by construction of $I_1$, $(\pi^*)^{-1}(I_1) \subset I_1 \cup I_2$. This implies that

$$\prod_{i\in I_1} \sigma^{\#}_{\pi(i)} = \prod_{j\in(\pi^*)^{-1}(I_1)} \sigma_j = \frac{\prod_{j\in I_1\cup I_2} \sigma_j}{\prod_{j\in(I_1\cup I_2)\setminus(\pi^*)^{-1}(I_1)} \sigma_j}. \tag{33}$$

Note also that the cardinality of the set $J_1 = (\pi^*)^{-1}(I_1)$ is equal to the cardinality of $I_1$, which implies that $|(I_1 \cup I_2) \setminus J_1| = |I_2|$. Combining (32), (33), and the last equality of cardinalities, we get

$$\prod_{\substack{i\in[n]\\ \pi(i)\neq\pi^*(i)}} \frac{2\sigma_i^2}{\sigma^2_{i,\pi(i)}} \le \frac{\prod_{j\in(I_1\cup I_2)\setminus J_1} \sigma_j}{\prod_{i\in I_2} \sigma^{\#}_{\pi(i)}} \le r_\sigma^{|I_2|}. \tag{34}$$

Using the same notation $I_1$ and $I_2$, we can check that

$$\kappa_{i,\pi(i)} \ge \begin{cases} \bar\kappa_{\text{in-in}}, & i \in I_1,\\ \bar\kappa_{\text{in-out}}, & i\in I_2.\end{cases}$$

Injecting this inequality into (31), and using (34), we get

$$\prod_{i=1}^n \frac{\|X_i - X^{\#}_{\pi^*(i)}\|^2}{\|X_i - X^{\#}_{\pi(i)}\|^2} \le \frac{r_\sigma^{I_2}\{2(d+\sqrt{d}\,\zeta_2)\}^{|I_1|+|I_2|}}{(\bar\kappa^2_{\text{in-in}} + 2d - 2\sqrt{d}\,\zeta_2)_+^{I_1}(\bar\kappa^2_{\text{in-out}} + 2d - 2\sqrt{d}\,\zeta_2)_+^{I_2}}.$$

Recall that this inequality is true on the event $\zeta_1 \le \bar\kappa/4$. It follows from last display that as soon as

$$\begin{cases} \zeta_1 & \le \bar\kappa/4\\ 4\sqrt{d}\,\zeta_2 & < \bar\kappa^2_{\text{in-in}}\\ 2d(r_\sigma - 1) + 4r_\sigma\sqrt{d}\,\zeta_2 & \le \bar\kappa^2_{\text{in-out}}\end{cases} \tag{35}$$

we have

$$\prod_{i=1}^n \frac{\|X_i - X^{\#}_{\pi^*(i)}\|^2}{\|X_i - X^{\#}_{\pi(i)}\|^2} < 1$$

for every $\pi$. It remains to show that, under the conditions of Theorem 4, the event in (35) has a probability at least $1-\alpha$. This will be done by using tail bounds for Gaussian and khi-squared distributions, combined with the union bound.

On the one hand, using the well-known tail bound for the standard Gaussian distribution and the union bound, we get

$$\mathbf{P}\bigg(\zeta_1 \ge \sqrt{2\log\Big(\frac{4nm}{\alpha}\Big)}\bigg) \le \sum_{i\neq j}\mathbf{P}\bigg(|\zeta_{i,j}| \ge \sqrt{2\log\Big(\frac{4nm}{\alpha}\Big)}\bigg) \le \alpha/2.$$

On the other hand, Lemma 1 and the union bound entail

$$\mathbf{P}\bigg(\zeta_2 \ge 2\sqrt{\log(4nm/\alpha)} + \frac{2\log(4nm/\alpha)}{\sqrt{d}}\bigg) \le \alpha/2.$$

Therefore, if

$$\begin{cases} \bar\kappa & \ge\ 4\sqrt{2\log(4nm/\alpha)}\\ \bar\kappa^2_{\text{in-in}} & \ge\ 8\sqrt{d\log(4nm/\alpha)} + 8\log(4nm/\alpha)\\ \bar\kappa^2_{\text{in-out}} & \ge\ 2d(r_\sigma-1) + 8r_\sigma\sqrt{d\log(4nm/\alpha)}\ + 8r_\sigma\log(4nm/\alpha)\end{cases}$$

then, on an event of probability $\ge 1-\alpha$, all the inequalities in (35) hold true. This completes the proof of the theorem.

# C General lower bound

In this section we formulate and prove a lower bound over all injective mappings $\pi : [n] \to [m]$. The theorem states that the rate presented and proved in Theorem 1 is indeed optimal. We show that even if $\bar{\kappa}_{\text{in-in}}$ and $\bar{\kappa}_{\text{in-out}}$ are of order $(d\log(nm))^{1/4} \vee (\log(nm))^{1/2}$ there are indeed scenarios in which any estimator $\hat{\pi}$ fails to detect $\pi^*$ with probability at least $1/3$.

**Theorem 5 (General lower bound)** *Denote $\kappa = \min\{\bar{\kappa}_{\text{in-in}}, \bar{\kappa}_{\text{in-out}}\}$. Assume that $m > n \geq 5$ and $d \geq 16\log(nm)$. Then, there exists a triplet $(\boldsymbol{\sigma}^\#, \boldsymbol{\theta}^\#, \pi^*)$ such that $6\kappa \geq (d\log(nm))^{1/4}$ and*

$$\inf_{\hat{\pi}} \mathbf{P}_{\boldsymbol{\theta}^\#, \boldsymbol{\sigma}^\#, \pi^*}(\hat{\pi} \neq \pi^*) > 1/3,$$

*where the infimum is taken over all injective matching maps $\pi : [n] \to [m]$.*

**Proof** We denote the set of all injective functions $\pi : [n] \to [m]$ as $\mathfrak{I}_{n,m}$. We use the notation $D(\mathbf{P}, \mathbf{Q})$ for the Kullback-Leibler (KL) divergence between two probability measures $\mathbf{P}$ and $\mathbf{Q}$ such that $\mathbf{P}$ is absolutely continuous with respect to $\mathbf{Q}$, $\mathbf{P} \ll \mathbf{Q}$. The identity mapping denoted by *id* is defined as follows: $id(i) = i,\ \forall i \in [n]$. It is also assumed that $\pi^* = id$.

To establish the general lower bound we use the following lemma:

**Lemma 3 (Tsybakov (2009), Theorem 2.5)** *Assume that for some integer $M \geq 2$ there exist distinct injective functions $\pi_0, \ldots, \pi_M \in \mathfrak{I}_{n,m}$ and mutually absolutely continuous probability measures $\mathbf{Q}_0, \ldots, \mathbf{Q}_M$ defined on a common probability space $(\mathcal{Z}, \mathscr{Z})$ such that*

$$\frac{1}{M}\sum_{j=1}^{M} D(\mathbf{Q}_j, \mathbf{Q}_0) \leq \frac{1}{8}\log M.$$

*Then, for every measurable mapping $\tilde{\pi} : \mathcal{Z} \to \mathfrak{I}_{n,m}$,*

$$\max_{j=0,\ldots,M} \mathbf{Q}_j(\tilde{\pi} \neq \pi_j) \geq \frac{\sqrt{M}}{\sqrt{M}+1}\Big(\frac{3}{4} - \frac{1}{2\sqrt{\log(M)}}\Big).$$

Since $d \geq 16\log(nm)$ then the rate from Theorem 1 becomes of order $(d\log(nm))^{1/4}$. We show that for $6\kappa \geq (d\log(nm))^{1/4}$ there is indeed a setting where the detection of $\pi^*$ fails with probability at least $1/4$ for any matching map $\tilde{\pi} \in \mathfrak{I}_{n,m}$. To show this we use Lemma 3 with properly chosen family of probability measures described in the following lemma.

**Lemma 4 (Collier and Dalalyan (2016), Lemma 14)** *Let $\varepsilon_1, \ldots, \varepsilon_m$ be real numbers defined by*

$$\varepsilon_k = \sqrt{2/d}\,\kappa\sigma_k^\#, \quad \forall k \in [m],$$

*and let $\mu$ be the uniform distribution on $\mathcal{E} = \{\pm\varepsilon_1\}^d \times \cdots \times \{\pm\varepsilon_m\}^d$. Denote by $\mathbf{P}_{\mu,\pi}$ the probability measure on $\mathbb{R}^{d\times m}$ defined by $\mathbf{P}_{\mu,\pi}(A) = \int_{\mathcal{E}} \mathbf{P}_{\boldsymbol{\theta},\pi}(A)\mu(d\boldsymbol{\theta})$. Let $\bar{\Theta}_\kappa$ be the set of $\boldsymbol{\theta}^\#$ such that $6\kappa \geq (d\log(nm))^{1/4}$. Assume that $\sigma_1^\# \leq \cdots \leq \sigma_m^\#$ and $\sigma_m^{\#2}/\sigma_1^{\#2} \leq 1 + \sqrt{\frac{\log(nm)}{16d}}$. Let $\pi = (k\ k')$ be the transposition that only permutes $k^{\text{th}}$ and $k'^{\text{th}}$ observations ($k < k'$). Then, the Kullback-Leibler divergence between $\mathbf{P}_{\mu,\pi}$ and $\mathbf{P}_{\mu,id}$ can be bounded as follows*

$$D(\mathbf{P}_{\mu,\pi}, \mathbf{P}_{\mu,id}) \leq \frac{1}{8}\log(m(m-1)/2).$$

*Additionally, $\mu(\mathcal{E} \setminus \bar{\Theta}_\kappa) \leq (m(m-1)/2)e^{-d/8}$.*

Applying Lemma 3 with $M = m(m-1)/2$, $\mathbf{Q}_0 = \mathbf{P}_{\mu,id}$ and $\{\mathbf{Q}_j\}_{j=1,\ldots,M} = \{\mathbf{P}_{\mu,\pi_{k,k'}}\}_{k\neq k'}$ we obtain that

$$\begin{aligned}
\inf_{\hat{\pi}} \max_{\pi^* \in \mathfrak{I}_{n,m}} \sup_{\boldsymbol{\theta}^\# \in \bar{\Theta}_\kappa} \mathbf{P}_{\boldsymbol{\theta}^\#, \boldsymbol{\sigma}^\#, \pi^*}(\hat{\pi} \neq \pi^*) &\geq \max_{\pi^* \in \{id\}\cup\{\pi_{k,k'}\}} \int_{\bar{\Theta}_\kappa} \mathbf{P}_{\boldsymbol{\theta}^\#, \pi^*}(\hat{\pi} \neq \pi^*)\frac{\mu(d\boldsymbol{\theta}^\#)}{\mu(\bar{\Theta}_\kappa)} \\
&\geq \max_{\pi^* \in \{id\}\cup\{\pi_{k,k'}\}} \mathbf{P}_{\mu,\pi^*}(\hat{\pi} \neq \pi^*) - \mu(\mathcal{E} \setminus \bar{\Theta}_\kappa) \\
&\geq \frac{\sqrt{15}}{\sqrt{15}+1}\Big(\frac{3}{4} - \frac{1}{2\sqrt{\log 15}}\Big) - \frac{m(m-1)}{2}e^{-d/8},
\end{aligned}$$

where the in the last inequality we applied the result of Lemma 3 in conjunction with the monotonicity of function $m \mapsto \frac{\sqrt{m}}{1+\sqrt{m}}(3/4 - (2\sqrt{\log(m)})^{-1})$. Recall that $m > n \geq 5$ and $d \geq 16\log(nm)$ yielding $\inf_{\hat{\pi}} \mathbf{P}_{\boldsymbol{\theta}^\sharp,\boldsymbol{\sigma}^\sharp,\pi^*}(\hat{\pi} \neq \pi^*) > 0.338$.

# D Further details on Experiment 3

In this section we present further details on real-data experiment presented in the paper. We first plot the estimation accuracy measured in the Hamming loss for two other scenes (Reichstag and Brandenburg Gate). The results are shown in Figure 5 in a similar manner as in Figure 4. We observe very similar behaviour in all 3 applied algorithms across scenes. From Figure 4 and Figure 5 we see that in general the image pairs from Reichstag scene are easier and LSL gets accuracy around $0.9$ when outlier rate, $(m-n)/n$, equals $70\%$. In the same situation for Brandenburg Gate scene the LSL accuracy is around $0.7$. This is due to quality, angle of the camera and other external factors of images.

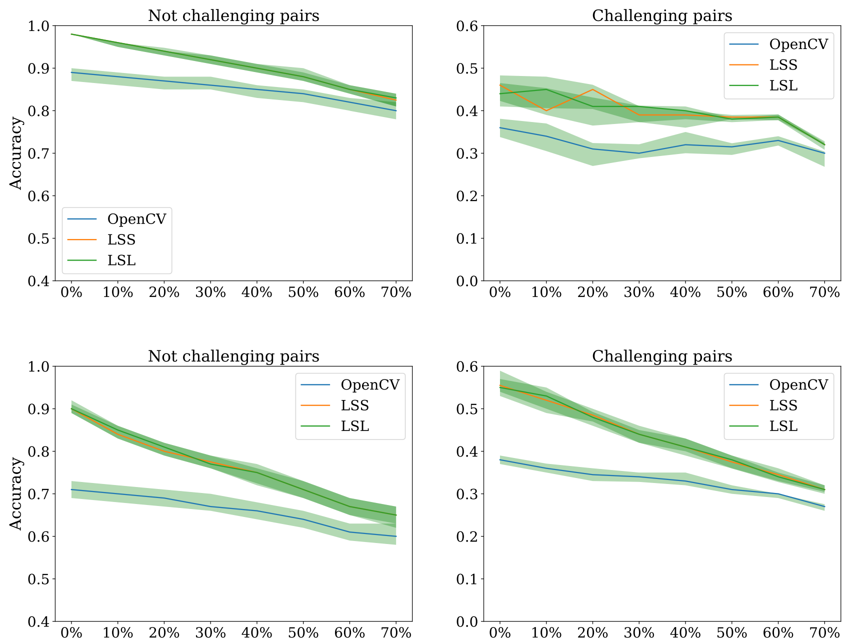


Figure 5: The estimation accuracy measured in the Hamming loss of the estimated matching in Exp. 3 for different values of the outlier rate, $(m-n)/n$, varying from $0\%$ to $70\%$. The medians of estimation accuracy both for challenging pairs (right plots) and simple pairs (left plots) of images from Reichstag (top) and Brandenburg Gate (bottom) scenes was computed using OpenCV, LSS and LSL matchers. The green region represents the interquartile range (lower and upper bounds being $25\%$ and $75\%$ percentiles, respectively).

We also plot the boxplots of the distances between SIFT descriptors of matching and non-matching keypoints both for easy (not challenging) and challenging pairs of images. Figure 6 has 3 plots for each of the scenes and 4 boxplots in each of them. The first 2 boxplots correspond to the distance between SIFT descriptors of matching and non-mathcing keypoints for easy pairs, while the last 2 boxplots are that of challenging pairs. There are several observations that are observed across all scenes. First, the median distance for matching keypoint descriptors is much less than that of non-matching keypoint descriptors. Second, the median distance between the matching keypoint descriptors from challenging pairs is much higher than that of easy pairs. We also observe that the distance distribution of non-matching keypoint descriptors is roughly the same for easy and challenging pairs. These observations suggests that distance based matching algorithms can be effectively applied.

To give a glimpse of what easy and challenging pairs of images look like we include sample pairs with accuracy of OpenCV matching algorithm greater than $0.5$ and sample pairs with accuracy less than $0.5$ from each scene. These pairs are collected in Figure 7.

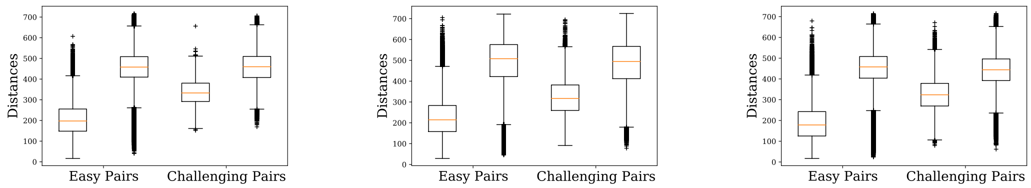


Figure 6: Boxplots of distances between SIFT descriptors for (left to right) Reichstag, Brandenburg Gate and Temple Nara scenes. We split datasets into easy and challenging pairs according to OpenCV matching algorithm score (the pairs with less than $0.5$ are considered challenging pairs, others are easy pairs). For each scene we then draw the boxplots of distances between descriptors of matching keypoints and non-matching keypoints grouped by easy and challenging pairs, respectively.

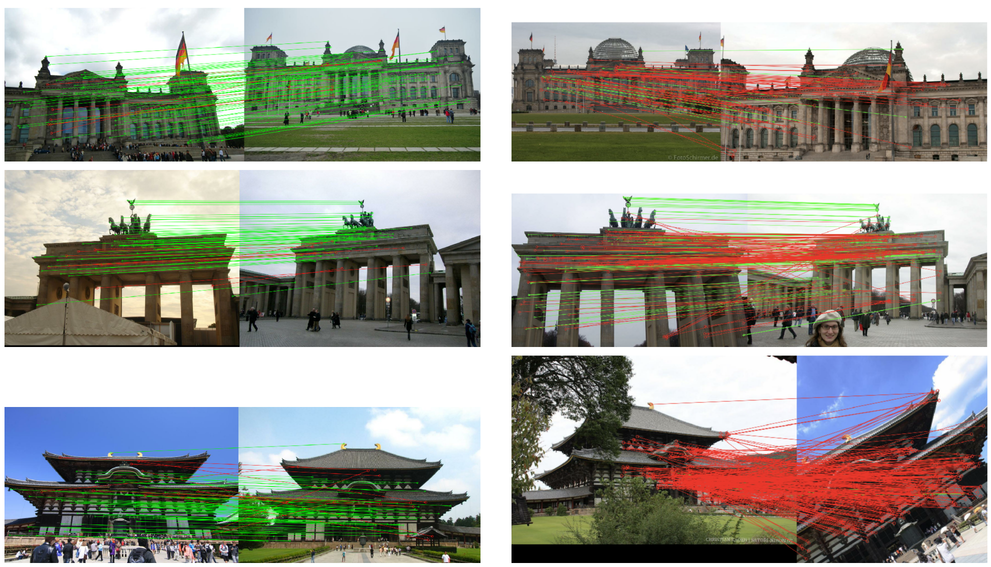

Figure 7: Matching map computed by LSL on randomly chosen easy (not challenging) and challenging pairs of images from each scene. The green lines represent the correct matching, and red lines are incorrect ones.