A Robust Method to Count and Locate Audio Sources in a Stereophonic Linear Instantaneous Mixture

Simon Arberet, Rémi Gribonval, and Frédéric Bimbot

IRISA, France

Abstract. We propose a robust method to estimate the number of audio sources and the mixing matrix in a linear instantaneous mixture, even with more sources than sensors. Our method is based on a multiscale Short Time Fourier Transform (STFT), and relies on the assumption that in the neighborhood of some (unknown) scales and time-frequency points, only one source contributes to the mixture. Such time-frequency regions provide local estimates of the corresponding columns of the mixing matrix. Our main contribution is a new clustering algorithm called DEMIX to estimate the number of sources and the mixing matrix based on such local estimates. In contrast to DUET or other similar sparsity-based algorithms, which rely on a global scatter plot, our algorithm exploits a local confidence measure to weight the influence of each time-frequency point in the estimated matrix. Inspired by the work of Deville, the confidence measure relies on the time-frequency local persistence of the activity/inactivity of each source. Experiments are provided with stereophonic mixtures and show the improved performance of DEMIX compared to K-means or ELBG clustering algorithms.

1 Introduction

The problem of estimating the number of audio sources and the mixing matrix is considered in a possibly degenerate noisy linear instantaneous mixture $x_m(\tau) = \sum_{n=1}^{N} a_{mn}s_n(\tau) + e_m(\tau)$, $1 \leq m \leq M$, more conveniently written in matrix form $\mathbf{x}(\tau) = \mathbf{A}\mathbf{s}(\tau) + \mathbf{e}(\tau)$. While the M signals $x_m(\tau)$ are observed, the number N of sources as well as the $M \times N$ mixing matrix $\mathbf{A}$, the N source signals $s_n(\tau)$ and the noise signals $e_m(\tau)$ are unknown.

Our approach relies on assumptions similar to those of DUET [1] and TIFROM [2,3]. It exploits the fact that for each source, there is at least one time-frequency region where it is the only source contributing to the mixture. This assumption is related to sparsity of the time-frequency representation of the sources, which is a well-known property of a variety of audio sources. In many sparsity-based source separation approaches [4,5,1] this property is exploited globally by drawing a scatter plot of the time-frequency values $\mathbf{X}(t, f)\}_{t,f}$ – which more or less displays lines directed by the columns $\mathbf{a}_n$ of the mixing matrix – and cluster them into N clusters. Such a global clustering approach is sensitive to the parameters of the clustering algorithm, and to the fact that the direction of some sources of weak energy might not appear clearly in the global scatter plot. Rather than using a full scatter plot, our approach is to exploit the local time-frequency persistence [2,3] of the activity/inactivity of each source to get a robust estimation of the number N of sources and the mixing matrix $\mathbf{A}$. This is similar to the TIFROM [2,3] method,

J. Rosca et al. (Eds.): ICA 2006, LNCS 3889, pp. 536–543, 2006.

which –in the stereophonic case– uses the variance of the ratio $\frac{X_2(t,f)}{X_1(t,f)}$ within a time-frequency region to determine whether the region contains a single active source or more. Our main contributions are to:

1. use a multi-resolution framework (multiple window STFT) to account for the different possible durations of audio structures in each source.
2. rely on a local confidence measure to determine how valid is the assumption that only one source contributes to the mixture in a given time-frequency region;
3. propose a new clustering algorithm called DEMIX, based on the confidence measure, that counts the sources and locates them.

In Section 2, after some reminders on related approaches to estimate the mixing matrix, we give the outline of our approach and describe the confidence measure. In Section 3 we describe the new clustering algorithm DEMIX, and Section 4 is devoted to experiments that compare several methods on audio mixtures.

2 Exploiting Sparsity and Persistence

Let us analyze briefly the most simple sparse source model: assume that at each time τ, only one source $n := n(\tau)$ is active ($s_n(\tau) \neq 0$ and $s_k(\tau) = 0\ \forall k \neq n$). In such a case, the noiseless mixture at time τ is $\mathbf{x}(\tau) = \mathbf{a}_n s_n(\tau)$. In other word each point $\mathbf{x}(\tau) \in \mathbb{R}^M$ is aligned on one of the columns $\mathbf{a}_n$ of the mixing matrix $\mathbf{A}$. In fact this simple model is not very sparse, but (the real and imaginary parts of) STFT values $\mathbf{X}(t,f)$ approximately displays such a behaviour, since the linear mixture model $\mathbf{X}(t,f) = \mathbf{A}\mathbf{S}(t,f) + \mathbf{E}(t,f)$ holds and in *many* time-frequency points (t,f), only one source is dominant compared to the others. However, there are points where several sources are similarly active, which can make it difficult to estimate the mixing matrix by simply clustering the global scatter plot.

2.1 Related Work

Many source separation methods for the stereophonic case ($M = 2$) use the idea of sparsity in order to find mixing directions. In Bofill and Zibulevsky's algorithm [4] and DUET [1], the global (time-frequency) scatter plot is transformed into angular values $\theta(t,f) = \tan^{-1}(X_2(t,f)/X_1(t,f))$, and the columns of the mixing matrix are estimated by finding maxima in an energy weighted smoothed histogram of these values. One of the difficulties with this approach is that it seems difficult to adjust how much smoothing must be performed on the histogram to resolve close directions without introducing spurious peaks.

Another approach is the TIFROM method [2,3] which consists in selecting only time-frequency points that have a great chance of being generated by only one source. In TIFROM, for each time-frequency point (t,f), the mean $\bar{\alpha}_{t,f}$ and variance $\sigma^2_{t,f}$ of Time-Frequency Ratios Of Mixtures $\alpha(t',f') = \widehat{x_2}(t',f')/\widehat{x_1}(t',f')$ are computed using all times t' within a neighborhood of t and $f' = f$. By searching for the lowest value of the variance, a time-frequency domain is located where essentially one source is present, and the corresponding column of $\mathbf{A}$ is identified as being proportional to $(1,\bar{\alpha}_{t,f})^T$.

However, it seems quite difficult to exploit TIFROM to actually determine *how many* sources are present in the mixture and find their directions. In addition, the *asymmetric* roles given by $\alpha(t',f')$ to the left and right channels of a stereophonic mixture is not fully satisfying as for sources located almost on the first channel (i.e., with mixing column close to $(0,1)^T$), the corresponding variance are likely to remain high, even at good time-frequency points.

2.2 Proposed Approach

We propose to overcome these limitations of TIFROM by replacing the local variance and mean of the ratios $\frac{\widehat{x_2}(t,f)}{\widehat{x_1}(t,f)}$ with the principal direction of the local scatter plot $(\widehat{x_1}(t,f),\widehat{x_2}(t,f))$, together with a measure of how strongly it points in its principal direction. For this, we first define time-frequency neighborhoods $\Omega_{t,f}$ around each time-frequency point (t,f). A discrete STFT with a window of size L computed with half overlapping windows and no zero padding provides values on the discrete time-frequency grid $t = kL/2$, $k \in \mathbb{Z}$ and $f = l/L$, $0 \leq l \leq L/2$. A possible shape of time-frequency neighborhood of a time-frequency point (t,f) is $\Omega_{t,f} = \{(t+kL/2, f+k'/L), |k| \leq S_T, |k'| \leq S_F\}$ but the approach is amenable to using or combining several shapes and size of neighborhoods. Each neighborhood provides a local scatter plot corresponding to a $M \times \mathrm{card}(\Omega_{t,f})$ matrix $\mathbf{X}_{\Omega_{t,f}}$ with entries $\mathrm{Re}[\mathbf{X}(t',f')]$ and $\mathrm{Im}[\mathbf{X}(t',f')]$ for $(t',f') \in \Omega_{t,f}$. Performing a Principal Component Analysis (PCA) on $\mathbf{X}_{\Omega_{t,f}}$ we obtain a principal direction as a unit vector $\hat{\mathbf{u}}(t,f) \in \mathbb{R}^M$. In the stereophonic case $M = 2$, the direction of the estimated principal unit vector $\hat{\mathbf{u}}(t,f) \in \mathbb{R}^2$ is equivalently translated into an angle $\hat{\theta}(t,f)$.

2.3 A Confidence Measure

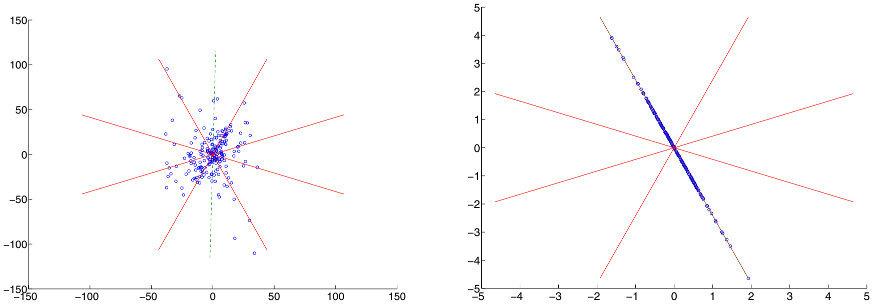
To have an idea of how likely it is that the unit principal vector $\hat{\mathbf{u}}(t,f)$ corresponds to a direction of the mixing matrix, we need to know with what confidence we can trust the fact that a single source is active in the corresponding local scatter plot. We propose to rely again on PCA to define the confidence measure

$$\widehat{\mathcal{T}}(t,f) := \hat{\lambda}_1(t,f) / \sum_{i=2}^{M} \hat{\lambda}_i(t,f) \tag{1}$$

where $\hat{\lambda}_1(t,f) \geq \ldots \geq \hat{\lambda}_M(t,f)$ are the eigenvalues of the $M \times M$ matrix $\mathbf{X}_{\Omega_{t,f}} \mathbf{X}_{\Omega_{t,f}}^T$. As explained in Appendix A, this measure can be viewed as a local signal to noise ratio between the dominant source and the contribution of the other ones together with the noise, so we will often express it in deciBels, that is to say $20 \log_{10} \widehat{\mathcal{T}}$.

Figure 1(a)-(b) shows the local scatter plot in two time-frequency regions: one where many sources are simultaneously active, and another one where essentially one source is active. It illustrates the good correlation of the value of the confidence measure with the validity of the tested hypothesis.

Figure 2(a) displays the collection of pairs $(\hat{\theta}(t,f), 20\log_{10} \hat{\mathcal{T}}(t,f))$, or *direction-confidence scatter plot* (DCSP), obtained by PCA for all time-frequency regions of the



(a) (b)

Fig. 1. Two local scatter plots for a stereophonic noiseless mixture of four audio sources. Solid lines indicate all possible true directions, the dashed line indicates the direction estimated by PCA. (a) Local scatter plot in a region where multiple sources contribute to the mixture. The measured confidence value is low (9.4 dB) (b) Region where essentially only one source contributes to the mixture. The measured confidence value is high (101.4 dB) and the dashed line coincides with one of the solid lines.

signal, together with four lines indicating the angles corresponding to the true underlying directions. One can observe that the higher the confidence, the smaller the average distance between the point and one of the true directions. We discuss in Appendix A a statistical analysis of the Significance of the confidence measure in the stereophonic case, which is used to build the DEMIX clustering algorithm described in the next section.

3 The DEMIX Algorithm

We propose a clustering algorithm called DEMIX (Direction Estimation of Mixing matrIX) which estimates both the number of sources and the directions of the columns of the mixing matrix. The algorithm is deterministic and does not rely on a prior knowledge on the number N of columns of $\mathbf{A}$. However, in the case where this number is known the algorithm can be adapted to incorporate this information. The algorithm is described in the stereophonic case $M = 2$ using angles $\hat{\theta}$ to denote mixing directions, but the approach extends to $M > 2$ mixtures by clustering the directions $\hat{\mathbf{u}}(t, f)$ instead.

The first step of the algorithm consists in iteratively creating K clusters by selecting points $(\widehat{\theta}_k, \widehat{\mathcal{T}_k})$ with highest confidence and aggregating sufficiently close points around them. The second step is to estimate the direction $\widehat{\theta}_k^c$ of each cluster. Finally, we use a statistical test to eliminate non significant clusters and keep $\widehat{N} \leq K$ clusters which centroids provide the estimated directions of the mixing matrix.

3.1 Step 1: Cluster Creation

DEMIX iteratively create K clusters $C_k \subset P$ –where P is the DCSP– starting from $K = 0$, $P_K = P_0 = P$:

1. find the point $(\widehat{\theta}_K, \widehat{\mathcal{T}_K}) \in P_K$ with the highest confidence;
2. create a cluster C_K with all points $(\widehat{\theta}, \widehat{\mathcal{T}}) \in P$ "sufficiently close" to $(\widehat{\theta}_K, \widehat{\mathcal{T}_K})$;
3. if $P_{K+1} := P_K \setminus C_K = \emptyset$, stop; otherwise increment $K \leftarrow K+1$ and go back to 1.

Note that in step 2 the newly created cluster might interesect previous clusters. To give a precise meaning to the notion of being "sufficiently close" to $(\widehat{\theta}_K, \widehat{\mathcal{T}_K})$, we rely on the statistical model developped in Appendix A and include in C_K all points $(\hat{\theta}, \widehat{\mathcal{T}})$ such that $|\hat{\theta} - \widehat{\theta}_K| \leq \sigma(\widehat{\mathcal{T}}, \widehat{\mathcal{T}_K})$ where the expression of $\sigma(\widehat{\mathcal{T}}, \widehat{\mathcal{T}_K})$ is given in Equation (8).

3.2 Step 2: Direction Estimation

Since the clusters might intersect, the estimation of the centroid $\widehat{\theta}_k^c$ of a cluster C_k is based on a subset $C_k'' \subset C_k$ of "unbiased" points that belong *exclusively* to C_k. Due to lack of space we skip the description of how these subsets are selected. In light of the statistical model developped in Appendix A, the points $(\hat{\theta}, \hat{\mathcal{T}}) \in C_k''$ are assumed independent and distributed as $\widehat{\theta} \sim \mathcal{N}\big(\theta_k^{true}, \sigma_\theta^2(\hat{\mathcal{T}})\big)$ where θ_k^{true} is the unknown underlying direction and $\sigma_\theta^2(\widehat{\mathcal{T}})$ is defined in equation (6). The centroid of the cluster if therefore defined as the minimum variance unbiased estimator of θ_k^{true}

$$\widehat{\theta}_k^c := \sum_{(\widehat{\theta}, \widehat{\mathcal{T}}) \in C_k''} \sigma_\theta^{-2}(\widehat{\mathcal{T}})\widehat{\theta} / \sum_{(\widehat{\theta}, \widehat{\mathcal{T}}) \in C_k''} \sigma_\theta^{-2}(\widehat{\mathcal{T}}). \tag{2}$$

3.3 Step 3: Cluster Elimination

The last step aims at removing possibly spurious clusters among the K that have been built. We propose to use the variance $1/\sum_{(\hat{\theta}, \hat{\mathcal{T}}) \in C_k''} \sigma_\theta^{-2}(\hat{\mathcal{T}})$ of the centroid estimator $\widehat{\theta}_k^c$ to help decide which clusters should be kept. We define two strategies: (DEMIXN) if we know the true number N of true directions, we keep the directions of the N clusters with the smallest centroid variance; (DEMIX) otherwise, we remove the directions of a clusters C_j whenever there is another cluster $C_o \neq C_j$ with

$$|\widehat{\theta}_j^c - \widehat{\theta}_o^c| \leq q_2 / \sum_{(\hat{\theta}, \hat{\mathcal{T}}) \in C_j''} \sigma_\theta^{-2}(\hat{\mathcal{T}}) \tag{3}$$

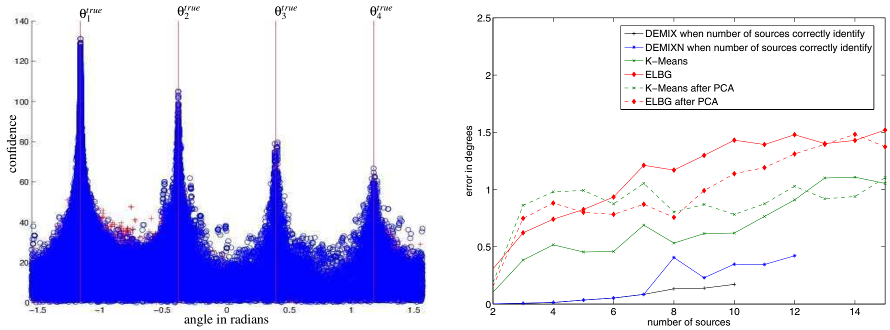
where the quantile q_2 defines a confidence interval (see the Appendix). It is also possible to replace σ_θ with a slightly modified version $\hat{\sigma}_\theta$ relying on a quantile q_1 to define a confidence interval, see Eq. (7). To finish, we recompute the centroids of the clusters defined by the remaining directions, as described in Sections 3.1 and 3.2.

4 Experiments

We compared on several test mixtures the proposed algorithms (DEMIX and DEMIXN) and the classical K-means [6] and ELBG [7] clustering algorithms. Two variants of

K-means and ELBG were considered, one on the scatter plot of $tan^{-1}(X_2/X_1)(t,f)$, the other one on that of the angles $\hat{\theta}(t,f)$ obtained after the proposed local PCA. The mixtures were based on signals taken from a set of 200 Polish voice excerpts of 5 seconds sampled at 4kHz[1]. Noiseless linear instantaneous mixtures were performed with mixing matrices in the most favorable shape where all directions are equally spaced (as in [4]), with a number of directions ranging from $N = 2$ to $N = 15$. For each N, we chose $T = 20$ differents configurations of signals sources among the 200 available. A first measure of performance was the rate of success in the estimation of the number of sources (for DEMX and DEMXN only, because K-means and ELBG have a fix number of clusters). We observed that up to $N = 8$ sources, DEMIX estimates correctly the number of directions in more than four cases out of five, but when $N > 10$ it always fails to count the number of sources. DEMIXN is similarly successful up to $N = 10$ sources and always fails for $N > 12$. The reason why DEMIXN can fail in finding the right number of sources while it is known is that the cluster creation stage might result in $K < N$ clusters. In case success, we could also measure the *angular mean error* (AME) which is the mean distance in degrees between true directions and estimated ones. Distances are computed in the best way to pair estimated directions with the true ones. For each tested algorithm, we computed the *average* AME among test mixtures where $\widehat{N} = N$. Since K-means and ELBG are randomly initialized, we ran them $I = 10$ times for each test mixture and focussed on the smallest AME over these 10 runs, which gives an optimistic estimate of their performance.

As can be seen on Figure 2(b), DEMIX and DEMIXN algorithms obtain the best performance. Since the AME for DEMIX and DEMIXN can only be measured when a correct number of sources is estimated, it is not computed when $N > 10$ (resp. $N > 12$) for DEMIX (resp. DEMIXN).



(a) Direction-confidence scatter plot (DCSP)

(b) Average AME as a function of the number of sources

Fig. 2. (a) Direction-confidence scatter plot of points $(\hat{\theta}, 20\log_{10}\hat{\mathcal{T}})$ obtained by PCA on time-frequency regions based on a single STFT with window size is $L = 4096$ and neighborhoods of size $|\Omega_{t,f}| = 10$. (see section 2.3). (b) Experimental results of section 4.

[1] The signals are available at http://mlsp2005.conwiz.dk/index.php?id=30

5 Conclusion

We designed,developped, and evaluated a new algorithm to estimate the source directions of the mixing matrix in the instantaneous underdetermined two-sensor case. The proposed DEMIX algorithm yields better experimental results than those obtained by K-means and ELBG clustering algorithms on the same multiscale STFT data. Furthermore DEMIX estimates itself the number of mixing sources. This algorithm was designed using a confidence measure which is one of the main contribution of the article. The confidence measure allows to well detect regions of time-frequency points where essentially one source is active. This confidence measure could also be used in the source separation process, in addition with the estimated mixing matrix, to determine which source should be estimated in which time-frequency region, possibly providing a fully adaptive local (pseudo) Wiener filter. Further works include the extension of the DEMIX algorithm to delayed and convolved mixtures. We are also looking into the practical aspects and validation of the algorithm for source separation with more than two sensors.

References

1. Yilmaz, O., Rickard, S.: Blind separation of speech mixtures via time-frequency masking. In: IEEE Transactions on Signal Processing. Volume 52. (2002) 1830–1847
2. F. Abrard, Y. Deville, P.W.: From blind source separation to blind source cancellation in the underdetermined case: a new approach based on time-frequency analysis. In: ICA. (2001)
3. F.Abrard, Y.: Blind separation of dependent sources using the "time-frequency ratio of mixtures" approach. In: ISSPA 2003, Paris, France, IEEE (2003)
4. P. Bofill, M.Z.: Underdetermined blind source separation using sparse representations. In: Signal Processing. Volume 81. (2001) 2353–2362
5. Paul D.O'Grady, B.A., T.Rickard, S.: Survey of sparse and non-sparse methods in source separation. IJIST (International Journal of Imaging Systems and Technology) (2005)
6. MacQueen, J.B.: Some methods for classification and analysis of multivariate observations. In: 5-th Berkeley Symposium on Mathematical Statistics and Probability. (1967)
7. Patanè, G., Russo, M.: The enhanced LBG algorithm. Neural Networks **14**(9) (2001) 1219–1237
8. Härdel, W., Simar, L., eds.: Applied multivariate statistical analysis. Spinger-Verlag (2003)

A Statistical Analysis in the Stereophonic Case

In this appendix we make a statistical model in the stereophonic case ($M = 2$) to better understand the significance of the confidence measure $\widehat{\mathcal{T}}(t,f)$ as a measure of how robustly $\widehat{\theta}(t,f)$ estimates the "true" underlying direction of the dominant source. For that, we model the STFT coefficients of the most active source in the time-frequency region $\Omega_{t,f}$ with a centered normal distribution of (large) variance σ^2, and the contribution of all other sources, plus possibly noise, as 2-dimensional centered normal distribution with covariance matrix $\widetilde{\sigma}^2\mathbf{Id}_2$. Letting $\mathbf{a}$ be the normalized ($\|\mathbf{a}\|^2 = 1$) column of the mixing matrix $\mathbf{A}$ which corresponds to the most active source, then the model is that for $(t',f') \in \Omega_{t,f}$ we have:

$$\mathbf{x}(t',f') = s(t',f')\mathbf{a} + \mathbf{n}(t',f') \tag{4}$$

where

$$s(t',f') \sim \mathcal{N}\left(0,\sigma^2\right),\ \mathbf{n}(t',f') \sim \mathcal{N}\left(0,\tilde{\sigma}^2\mathbf{Id}_2\right) \tag{5}$$

therefore $\mathbf{x}(t',f') \sim \mathcal{N}\left(0,\tilde{\sigma}^2\mathbf{Id}_2+\sigma^2\mathbf{aa}^T\right)$. Let $\lambda_1 \geq \lambda_2$ be the eigenvalues of the covariance matrix $\Sigma := \tilde{\sigma}^2\mathbf{Id}_2+\sigma^2\mathbf{aa}^T$ and $\mathbf{u}=(u_1,u_2)^T$ be a unit eigenvector corresponding with λ_1. By elementary linear algebra we have $\frac{\lambda_1}{\lambda_2}=\frac{\tilde{\sigma}^2+\sigma^2}{\tilde{\sigma}^2}=1+\frac{\sigma^2}{\tilde{\sigma}^2}$ and, if $\lambda_1>\lambda_2$ (i.e., $\sigma>0$), $\mathbf{u}$ is colinear to $\mathbf{a}$. Therefore, the true direction $\theta^{true}=\tan^{-1}(\frac{a_2}{a_1})$ is given by the direction of the principal component. Note that in this model λ_1/λ_2 is related to the "local signal to noise ratio" $\sigma^2/\tilde{\sigma}^2$ between the most active source and the others.

A.1 Precision of PCA

Since the values $\widehat{\theta}(t,f)$ and $\widehat{\mathcal{T}}(t,f)=\hat{\lambda}_1/\hat{\lambda}_2$ are computed by PCA on sample of $m :=$ card$(\Omega_{t,f})$ points, they only provide estimates of the true direction and of the "true" confidence λ_1/λ_2 with a finite precision which we want to estimate as a function of the sample size m. For that, we use the following result which is an immediate application of [8, Theorems 4.11, 5.7, 9.4] : for large sample size, $\widehat{\mathcal{T}}/(\lambda_1/\lambda_2)$ converges in law to $\mathcal{N}\left(1,\sigma^2_{\mathcal{T}}\right)$ with $\sigma^2_{\mathcal{T}}=4/(m-1)$, and $\widehat{\theta}$ converges in law to $\mathcal{N}(\theta^{true},\sigma^2_{\theta}(\lambda_1/\lambda_2))$ with

$$\sigma^2_{\theta}(\mathcal{T}) := \frac{1}{m-1}\frac{\mathcal{T}}{(\mathcal{T}-1)^2}. \tag{6}$$

A.2 Confidence Intervals

If λ_1/λ_2 is known, then we know the standard deviation of the estimated angle $\hat{\theta}$ with respect to the true one. Since we know the distribution of the confidence measure $\hat{\mathcal{T}}$ which is close, but not equal to λ_1/λ_2, we can only predict the deviation of $\hat{\theta}$ with respect to a "true" direction" using confidence intervals. With probability exceeding $1-\alpha(q_1)/2$, we have $\lambda_1/\lambda_2 \geq \hat{\mathcal{T}}/(1+q_1\sigma_{\mathcal{T}})$. Therefore, instead of $\sigma^2_{\theta}(\hat{\mathcal{T}})$ we can use

$$\hat{\sigma}^2_{\theta}(\hat{\mathcal{T}}) := \sigma^2_{\theta}\left(\hat{\mathcal{T}}/(1+q_1\sigma_{\mathcal{T}})\right) \tag{7}$$

and model $\hat{\theta}$ as $\hat{\theta} \sim \mathcal{N}\left(\theta^{true},\hat{\sigma}^2_{\theta}(\hat{\mathcal{T}})\right)$ instead of $\hat{\theta} \sim \mathcal{N}\left(\theta^{true},\sigma^2_{\theta}(\hat{\mathcal{T}})\right)$.

Neglecting the possible dependencies between $\hat{\theta}$ and $\hat{\mathcal{T}}$ and following the same path, we get a statistical *upper bound* $|\hat{\theta}-\theta^{true}| \leq q_2\hat{\sigma}_{\theta}(\hat{\mathcal{T}})$ with confidence level $1-\alpha(q_2)/2$. We use it to determine whether two points belong to the same cluster in the cluster creation step. This leads to the definition

$$\sigma(\widehat{\mathcal{T}},\widehat{\mathcal{T}}^c) = q_2\left(\hat{\sigma}_{\theta}(\widehat{\mathcal{T}})+\hat{\sigma}_{\theta}(\widehat{\mathcal{T}}^c)\right) \tag{8}$$

We use quantil values $q_1=q_2=2.33$ to provide confidence levels of 99 percent.