Algorithm 1 Robust Expectation-Maximization Algorithm for Simulated GMM

procedure EM_ALG_GMM(sampleMat, c, λ, d, τ, μ, Σ)
    $\mathbf{X} \leftarrow \text{sampleMat}_{n\times d}$
    $n \leftarrow$ number of total samples
    $d \leftarrow$ number of dimensions
    $\tau \leftarrow \tau_{c\times 1}$ ▷ Initial Guess of Proportion of Clusters
    $\Sigma \leftarrow \Sigma_{c\times d\times d}$ ▷ Initial Guess of covariance matrix of each cluster
    $\mu \leftarrow \mu_{c\times d}$ ▷ Initial Guess of the vector mean of each cluster
    $\mathbf{T} \leftarrow \mathbf{0}_{c\times n}$ ▷ This will contain the probability that each sample is in each cluster
    $\text{iter}_{max} \leftarrow 20$ ▷ Setting the maximum loops for the algorithm

    for threshold = (in) **do**
        $\ell \leftarrow \lambda^{\text{threshold}}$
        $\mu_{\text{old}} \leftarrow \mu$
        for iter $= 1 : \text{iter}_{max}$ **do**
            $\mathbf{T}_{old} \leftarrow \mathbf{T}$ ▷ So that we can break if the T plateaus
            ▷ **E Step: Find the posterior probabilities**
            for $j = 1 : c$ **do**
                $\mathbf{T}[j,] \leftarrow \log(\tau_j + \exp(\frac{-\frac{1}{2}(x-\mu_j)^T\Sigma^{-1}(x-\mu_j)}{\sqrt{(2\pi)^d\|\Sigma_j\|}}))$ ▷ log likelihood of points in a selected cluster
            end for
            for $i = 1 : n$ **do**
                scale $\leftarrow \max(\mathbf{T}[,i])$
                for $k = 1 : c$ **do**
                    $\mathbf{T}[k,i] \leftarrow \frac{\exp(\mathbf{T}[k,i] - \text{scale})}{\sum_{m=1}^{c}\exp(\mathbf{T}[m,i] - \text{scale})}$
                end for
            end for
            ▷ **M Step: update τ, μ, and Σ**
            for $j = 1 : c$ **do**
                $\tau_j \leftarrow (1/n) \times \sum_{i=1}^{n} \mathbf{T}[j,i]$ ▷ Update τ
                $\mathbf{E} \leftarrow \mathbf{0}_{n\times d}$ ▷ This error matrix will be used to filter out outliers
                $\mu'_j \leftarrow \begin{bmatrix} -\mu_j- \\ -\mu_j- \\ \vdots \\ -\mu_j- \end{bmatrix}_{n\times d}$
                $\mathbf{A} \leftarrow \text{RowSum}\left((\mathbf{X}-\mu'_j)\Sigma_j^{-1} \odot (\mathbf{X}-\mu'_j)\right)$ ▷ Mahalanobis distance of each point to the cluster mean
                for $k = 1 : n$ **do**
                    if $A_k < \ell^2$ **then**
                        $E[k,] \leftarrow \mathbf{0}_{1\times d}$
                    else
                        $E[k,] \leftarrow X[k,] - \mu_j$
                    end if
                end for

$\mu_j \leftarrow \dfrac{\sum_{i=1}^{n} \mathbf{T}[j,i] \times (\mathbf{X}[i,] - \mathbf{E}[i,])}{\sum_{i=1}^{n} \mathbf{T}[j,i]}$ ▷ Update μ

$\mathbf{q}_{k\times 1} \leftarrow$ Row Indices for which the error term is zero ▷ Filter out those outliers when calculating the covariance matrix

if unstructure **then**

$\Sigma_j \leftarrow \dfrac{\sum_{i=1}^{k} \mathbf{T}[j,q_k] \times (\mathbf{X}[q_k,] - \mathbf{E}[q_k,] - \mu_j)^\top (\mathbf{X}[q_k,] - \mathbf{E}[q_k,] - \mu_j)}{\sum_{i=1}^{k} \mathbf{T}[j,q_k]}$ ▷ Update Σ

end if

if constant **then**

$\Sigma_j \leftarrow \sum \dfrac{\mathbf{T}[j,q_k] \times (\mathbf{X}[q_k,] - \mathbf{E}[q_k,] - \mu_j)^2}{d} \times \mathbf{I}_{d\times d}$

end if

if diagonal **then**

$\Sigma_j \leftarrow diag(\dfrac{\sum_{i=1}^{k} \mathbf{T}[j,q_k] \times (\mathbf{X}[q_k,] - \mathbf{E}[q_k,] - \mu_j)^\top (\mathbf{X}[q_k,] - \mathbf{E}[q_k,] - \mu_j)}{\sum_{i=1}^{k} \mathbf{T}[j,q_k]})$ ▷ Update Σ

end if

end for

if Number of Non-zero Elements in $\mathbf{T}[j,]$ smaller than 3 **then** ▷ Watch out for uninformative clusters

$\mu_j \leftarrow \mathbf{0}_{1\times d}$

$\Sigma_j \leftarrow 0.01 \times \mathbf{I}_{d\times d}$

end if

if $\max(|\mathbf{T}_{old} - \mathbf{T}|) < 1 \times 10^{-6}$ **then**

break

end if

end for

end for

end procedure