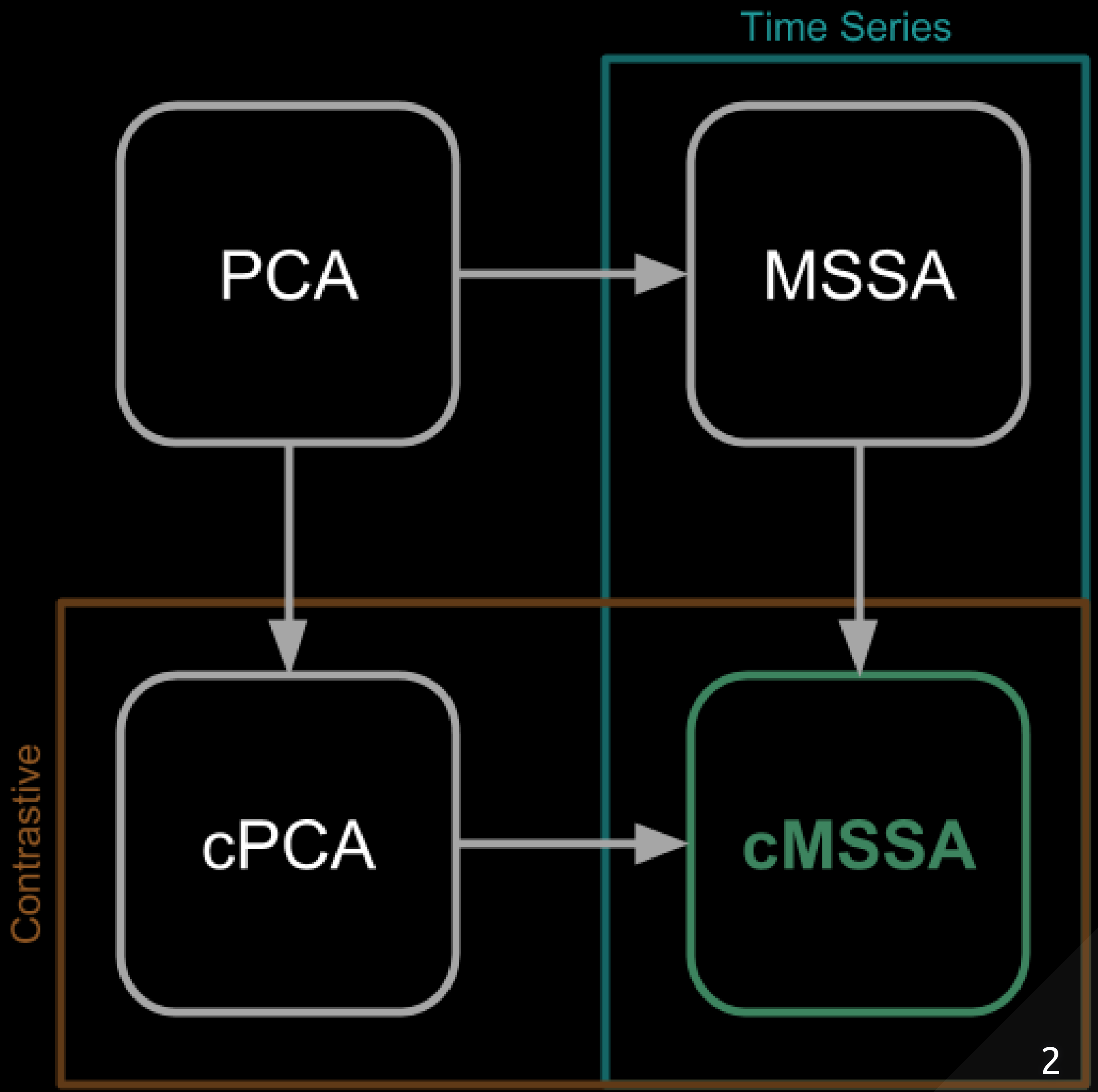
Contrastive Multivariate Singular Spectrum Analysis

Abdi-Hakin Dirie[†], Abubakar Abid[‡], James Zou[‡]

[†] Diffeo Labs, [‡] Stanford University

Marrying Two Ideas

- Multivariate Singular Spectrum Analysis (MSSA): A PCA-like method for unsupervised signal decomposition.
- cPCA [Abid 2018]: An extension to PCA that uses the idea of *contrastive learning*.



Algorithm

$$X \in \mathbb{R}^{T \times D}$$

A centered, weakly-stationary
D-channel time series for T steps.

Construct the prerequisite "Hankel" matrix:

$$H_{\mathbf{x}} = \begin{pmatrix} x_1 & x_2 & \cdots & x_W \\ x_2 & x_3 & \cdots & x_{W+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{T'} & x_{T'+1} & \cdots & x_T \end{pmatrix}$$

$$H_X = \left[H_{\mathbf{x}^{(1)}}; H_{\mathbf{x}^{(2)}}; \ldots; H_{\mathbf{x}^{(D)}} \right]$$

Compute the covariance matrix:

$$C_X = \frac{1}{T'} H_X^T H_X$$

(Repeat for a background signal Y to get C_Y.)

The "c" in cMSSA

$C = C_X - \alpha C_Y$ for $\alpha \geq 0$

C (hopefully) captures covariance information specific to X and not Y.

From here it's the same as standard MSSA:

- Find $E \in \mathbb{R}^{DW \times K}$, the top K eigenvectors of C.
- Principal component (PC) space: $A = H_X E$.
- Reconstructed component (RC) space:

$$R_{tj}^{(k)} = \frac{1}{W_t} \sum_{t'=L_t}^{U_t} A_{t-t'+1,k} \cdot \mathbf{e}_{(j-1)W+t'}^{(k)}$$

for component $k \in \{1, \ldots, K\}$

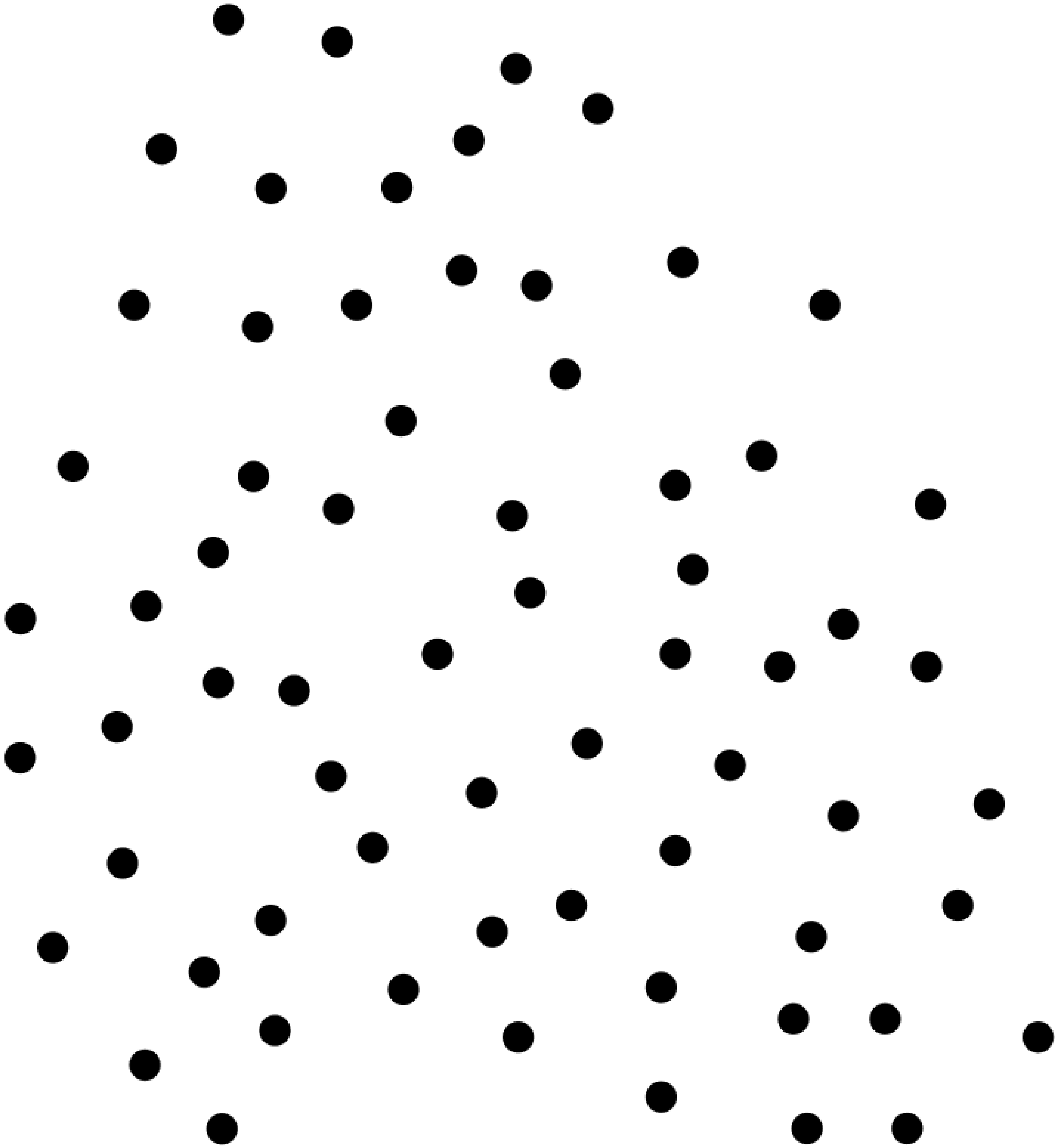
How do we determine α?

Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

procedure
- $Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$
- **for** $\alpha^{(i)} \in Q$ **do**
  - $H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$
  - $C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$
  - $E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$
- $S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$
- **for** $i \in \{1, \ldots, n+1\}, j \in \{i, \ldots, n+1\}$ **do**
  - $S_{i,j}, S_{j,i} \leftarrow \left\| {E^{(i)}}^T E^{(j)} \right\|_*$
- $Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$
- $Q^* \leftarrow \{0\}$
- **for** $z \in Z$ **do**
  - **if** $0 \notin z$ **then**
    - $\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$
    - $Q^* \leftarrow Q^* \cup \{\alpha^*\}$

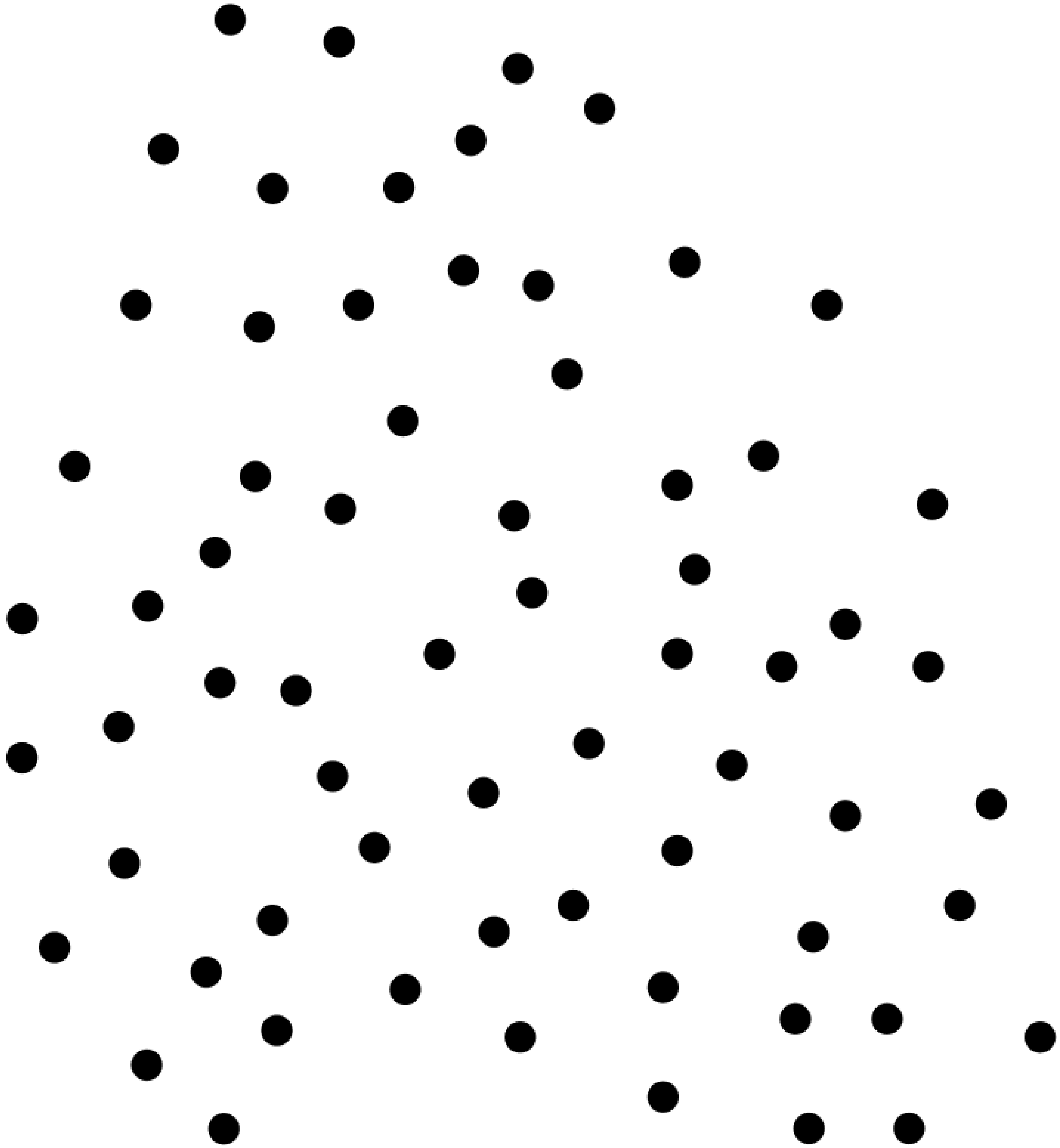
return Q^*, set of m best αs, including zero.



Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

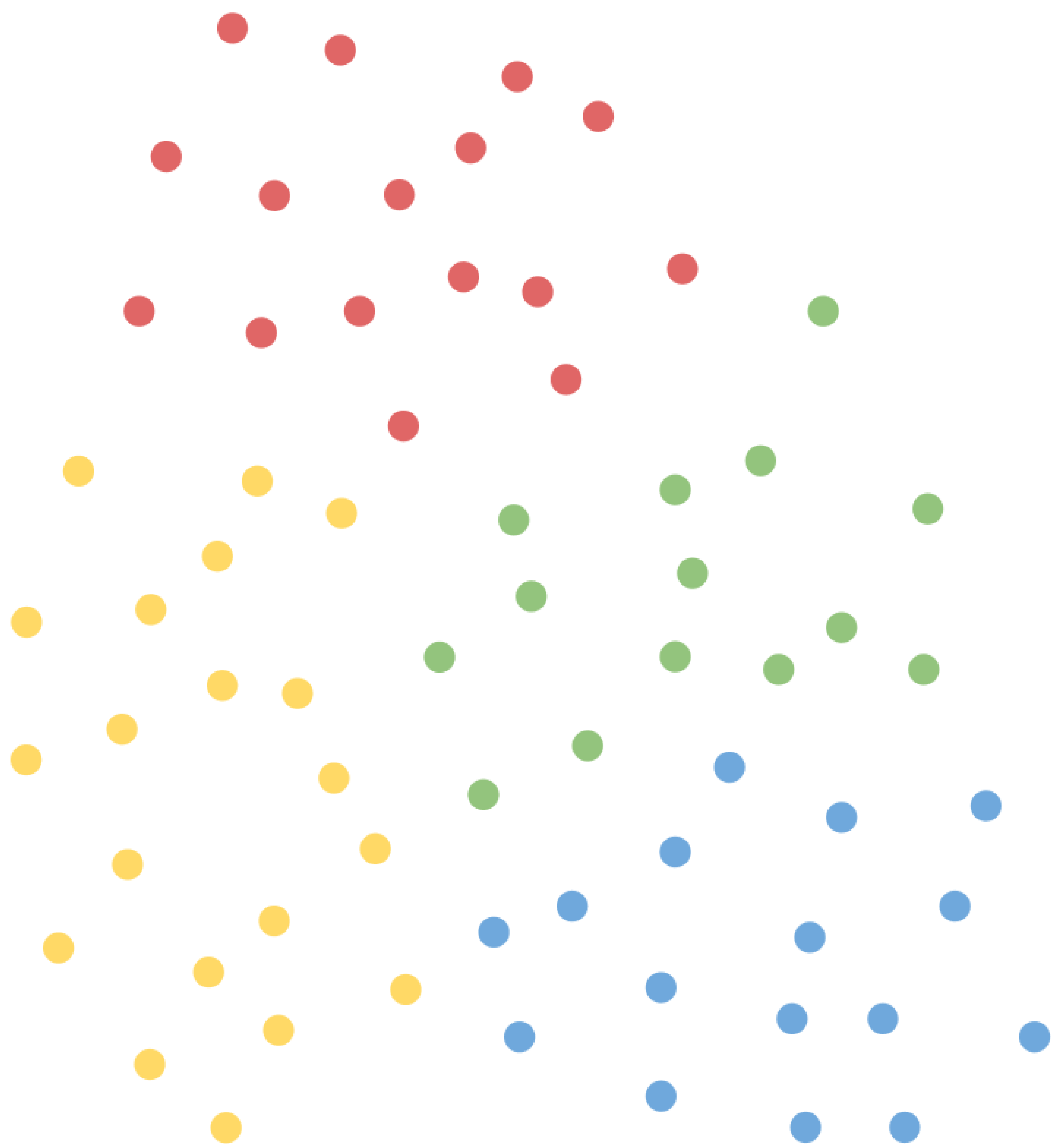
```
procedure
    Q ← LogSpace(α_min, α_max, n) ∪ {0}
    for α^(i) ∈ Q do
        H_X, H_Y ← Hankel(X, W), Hankel(Y, W)
        C_X, C_Y ← Cov(H_X), Cov(H_Y)
        E^(i) ← EigenDecomp(C_X − α^(i) C_Y, K)
    S ← Empty(ℝ^(n+1×n+1))
    for i ∈ {1, ..., n+1}, j ∈ {i, ..., n+1} do
        S_{i,j}, S_{j,i} ← ‖E^(i)T E^(j)‖_*
    Z ← SpectralCluster(S, Q, m)
    Q* ← {0}
    for z ∈ Z do
        if 0 ∉ z then
            α* ← ClusterMediod(z, S)
            Q* ← Q* ∪ {α*}
return Q*, set of m best αs, including zero.
```



Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

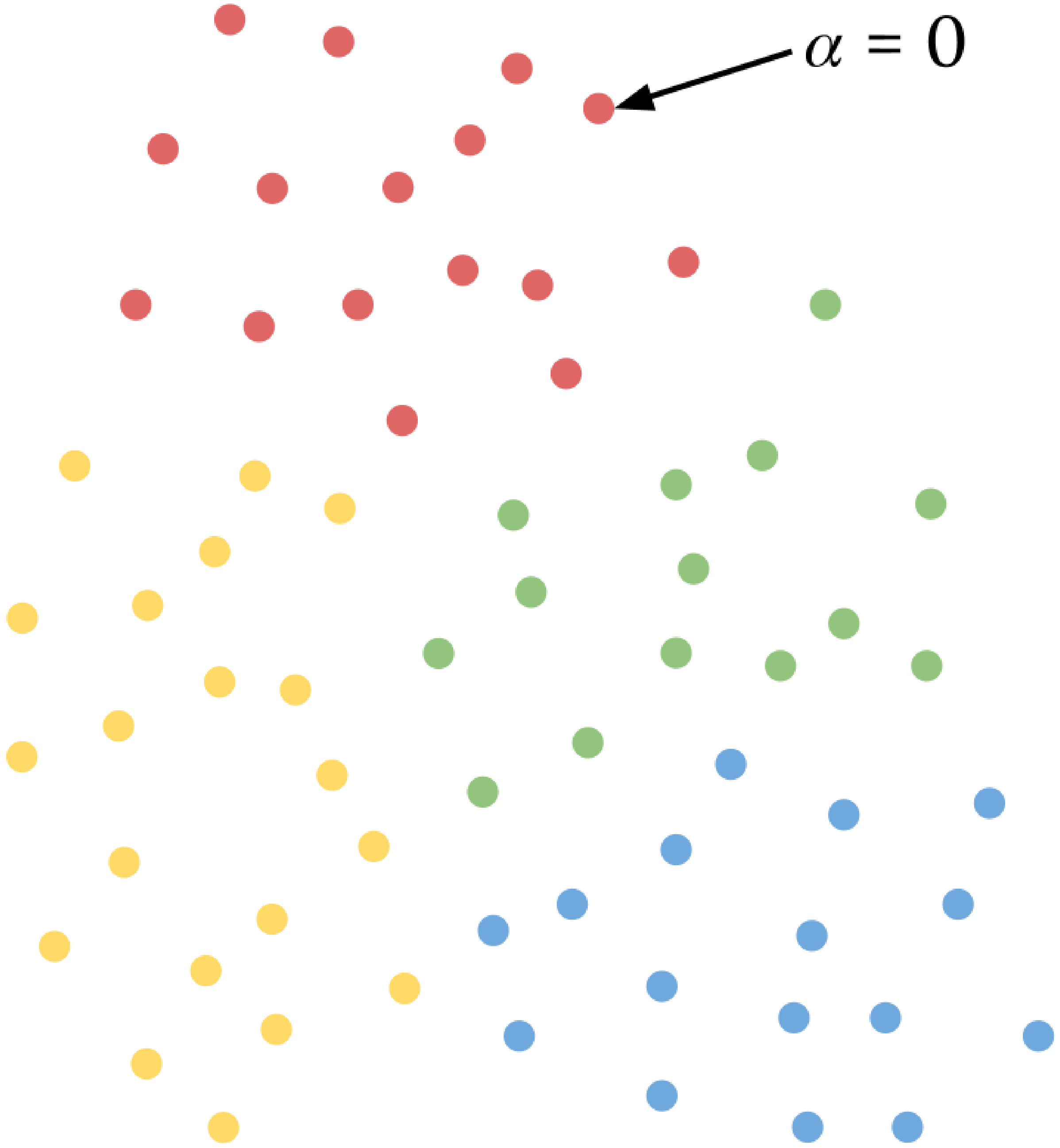
procedure
    $Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$
    for $\alpha^{(i)} \in Q$ **do**
        $H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$
        $C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$
        $E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$
    $S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$
    for $i \in \{1, \ldots, n+1\}, j \in \{i, \ldots, n+1\}$ **do**
        $S_{i,j}, S_{j,i} \leftarrow \left\| {E^{(i)}}^T E^{(j)} \right\|_*$
    $Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$
    $Q^* \leftarrow \{0\}$
    for $z \in Z$ **do**
        if $0 \notin z$ **then**
            $\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$
            $Q^* \leftarrow Q^* \cup \{\alpha^*\}$
return Q^*, set of m best αs, including zero.



Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

procedure
    $Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$
    for $\alpha^{(i)} \in Q$ **do**
        $H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$
        $C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$
        $E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$
    $S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$
    for $i \in \{1, \ldots, n+1\}, j \in \{i, \ldots, n+1\}$ **do**
        $S_{i,j}, S_{j,i} \leftarrow \left\| {E^{(i)}}^T E^{(j)} \right\|_*$
    $Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$
    $Q^* \leftarrow \{0\}$
    for $z \in Z$ **do**
        if $0 \notin z$ **then**
            $\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$
            $Q^* \leftarrow Q^* \cup \{\alpha^*\}$
return Q^*, set of m best αs, including zero.



Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

procedure

$Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$

for $\alpha^{(i)} \in Q$ **do**

$H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$

$C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$

$E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$

$S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$

for $i \in \{1, \ldots, n+1\}, j \in \{i, \ldots, n+1\}$ **do**

$S_{i,j}, S_{j,i} \leftarrow \left\| {E^{(i)}}^T E^{(j)} \right\|_*$

$Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$

$Q^* \leftarrow \{0\}$

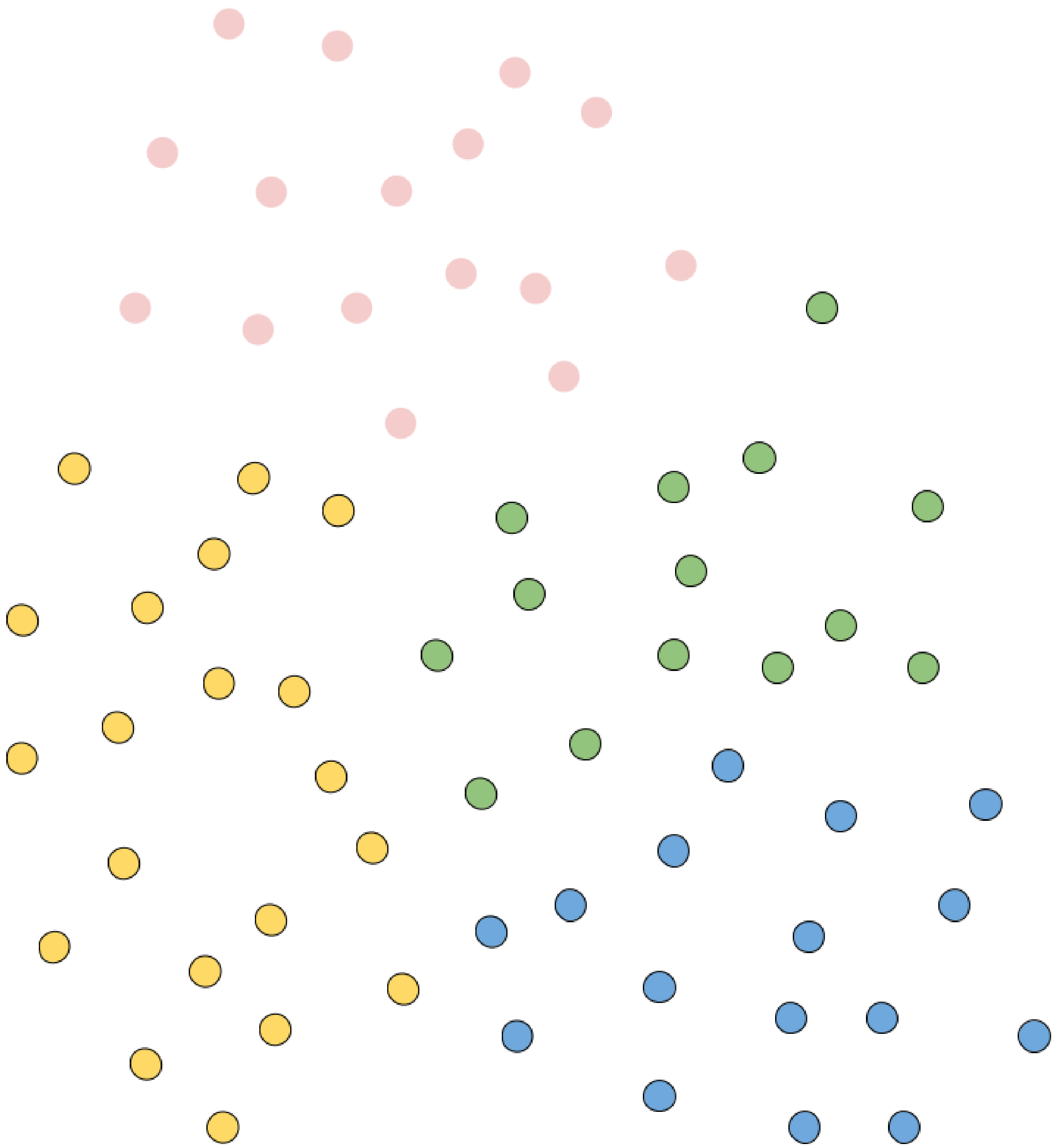
for $z \in Z$ **do**

if $0 \notin z$ **then**

$\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$

$Q^* \leftarrow Q^* \cup \{\alpha^*\}$

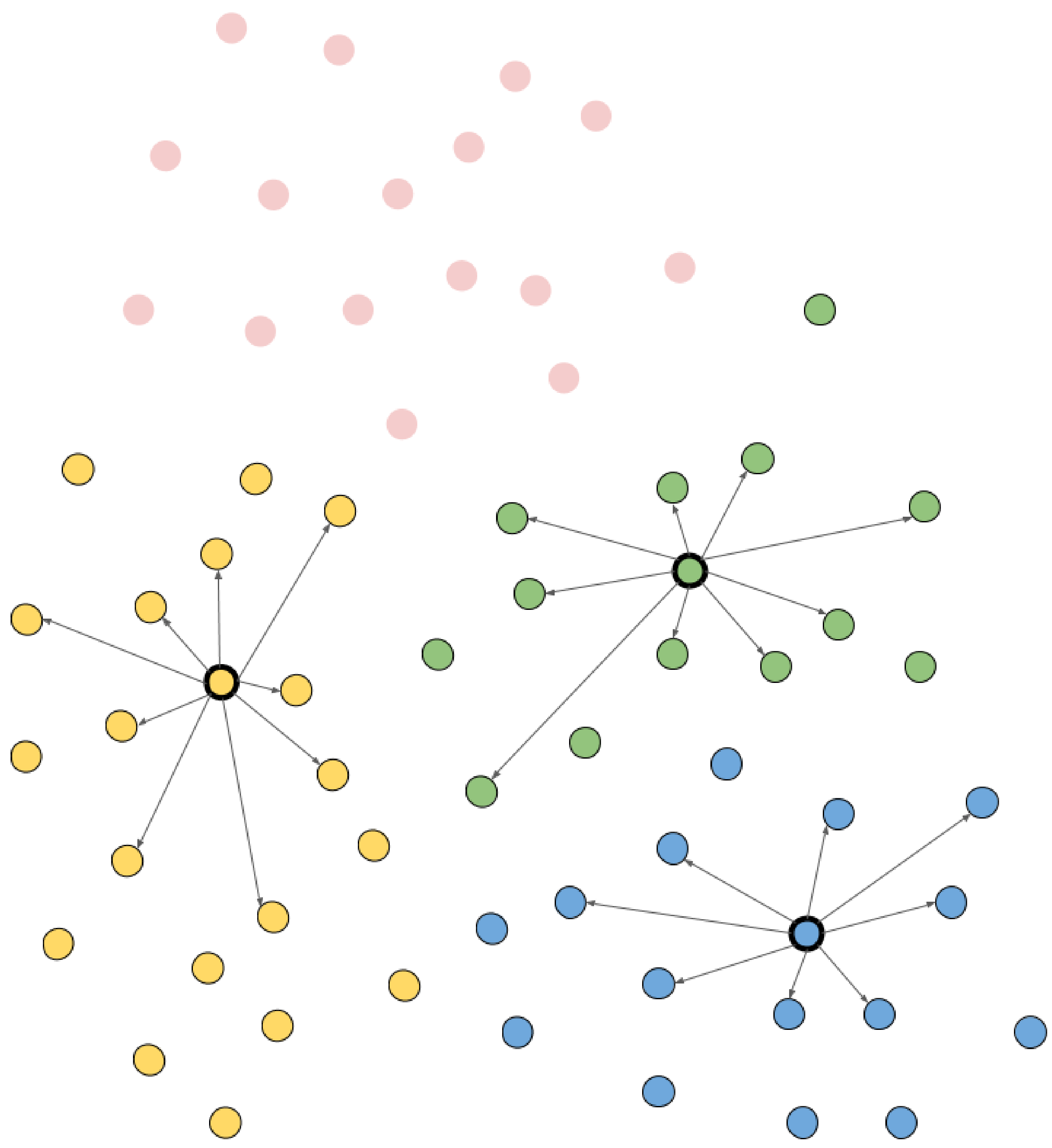
return Q^*, set of m best αs, including zero.



Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

```
procedure
    Q ← LogSpace(α_min, α_max, n) ∪ {0}
    for α^(i) ∈ Q do
        H_X, H_Y ← Hankel(X, W), Hankel(Y, W)
        C_X, C_Y ← Cov(H_X), Cov(H_Y)
        E^(i) ← EigenDecomp(C_X − α^(i) C_Y, K)
    S ← Empty(ℝ^(n+1×n+1))
    for i ∈ {1, . . . , n + 1}, j ∈ {i, . . . , n + 1} do
        S_{i,j}, S_{j,i} ← ‖E^(i)^T E^(j)‖_*
    Z ← SpectralCluster(S, Q, m)
    Q* ← {0}
    for z ∈ Z do
        if 0 ∉ z then
            α* ← ClusterMediod(z, S)
            Q* ← Q* ∪ {α*}
return Q*, set of m best αs, including zero.
```

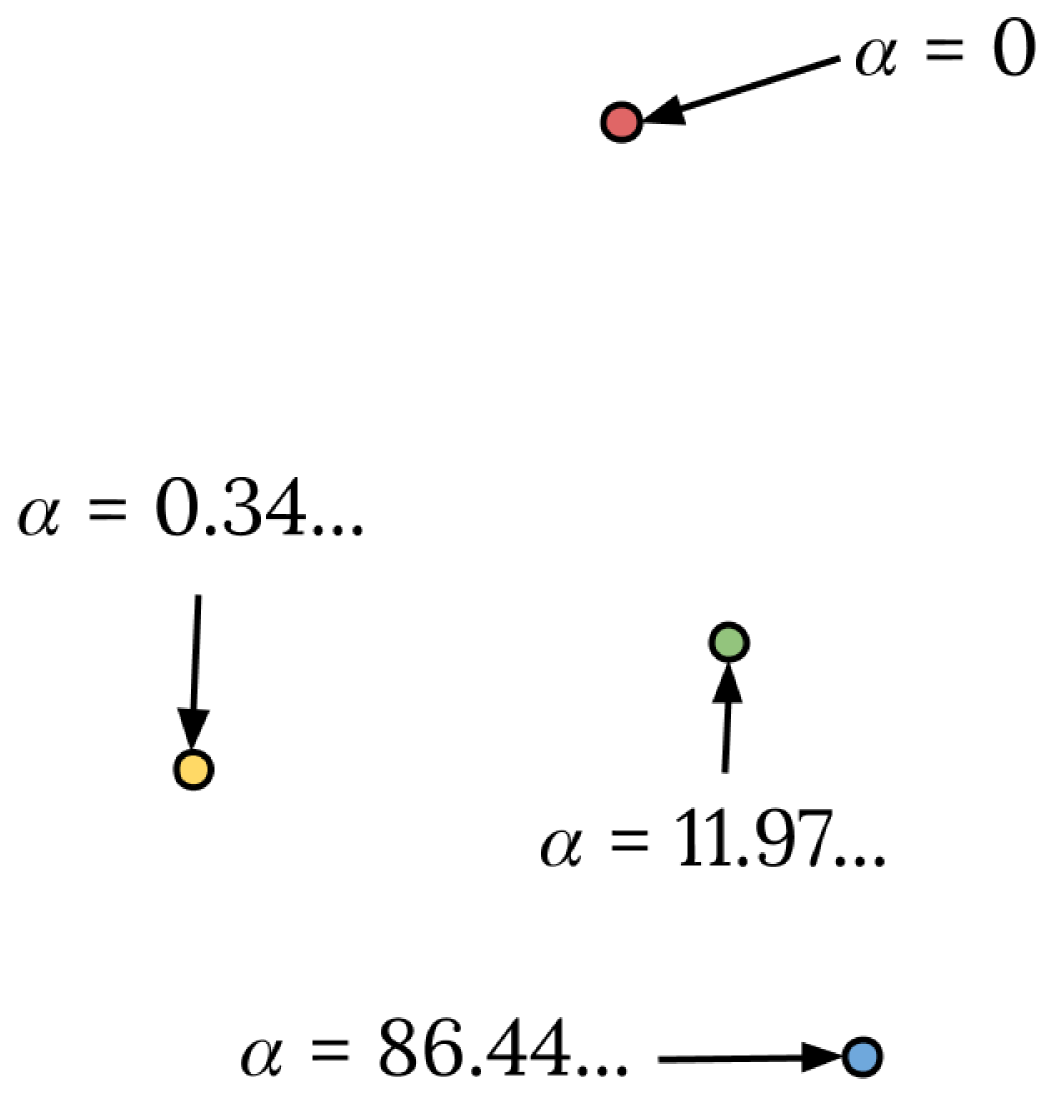


Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

procedure
- $Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$
- **for** $\alpha^{(i)} \in Q$ **do**
  - $H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$
  - $C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$
  - $E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$
- $S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$
- **for** $i \in \{1, \ldots, n+1\}$, $j \in \{i, \ldots, n+1\}$ **do**
  - $S_{i,j}, S_{j,i} \leftarrow \left\| E^{(i)^T} E^{(j)} \right\|_*$
- $Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$
- $Q^* \leftarrow \{0\}$
- **for** $z \in Z$ **do**
  - **if** $0 \notin z$ **then**
    - $\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$
    - $Q^* \leftarrow Q^* \cup \{\alpha^*\}$

return Q^*, set of m best αs, including zero.



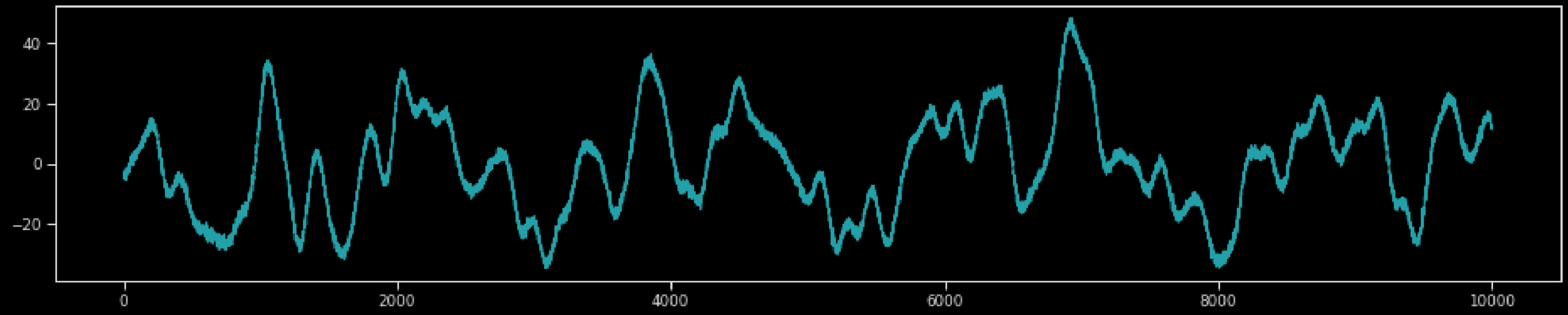
Algorithm 1 Spectral α-Search

Require: Minimum α to consider $\alpha_{\min}$, maximum α to consider $\alpha_{\max}$, number of αs to consider n, number of αs to return m, foreground signal X, background signal Y, window W, and number of components K.

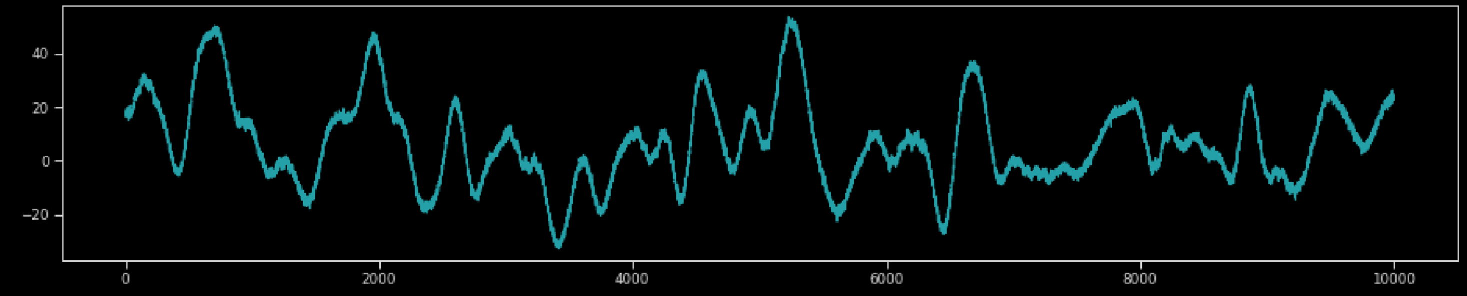
procedure
    $Q \leftarrow \textsc{LogSpace}(\alpha_{\min}, \alpha_{\max}, n) \cup \{0\}$
    for $\alpha^{(i)} \in Q$ **do**
        $H_X, H_Y \leftarrow \textsc{Hankel}(X, W), \textsc{Hankel}(Y, W)$
        $C_X, C_Y \leftarrow \textsc{Cov}(H_X), \textsc{Cov}(H_Y)$
        $E^{(i)} \leftarrow \textsc{EigenDecomp}(C_X - \alpha^{(i)} C_Y, K)$
    $S \leftarrow \textsc{Empty}(\mathbb{R}^{n+1 \times n+1})$
    for $i \in \{1, \ldots, n+1\}, j \in \{i, \ldots, n+1\}$ **do**
        $S_{i,j}, S_{j,i} \leftarrow \left\| {E^{(i)}}^T E^{(j)} \right\|_*$
    $Z \leftarrow \textsc{SpectralCluster}(S, Q, m)$
    $Q^* \leftarrow \{0\}$
    for $z \in Z$ **do**
        if $0 \notin z$ **then**
            $\alpha^* \leftarrow \textsc{ClusterMediod}(z, S)$
            $Q^* \leftarrow Q^* \cup \{\alpha^*\}$
return Q^*, set of m best αs, including zero.

Synthetic Example

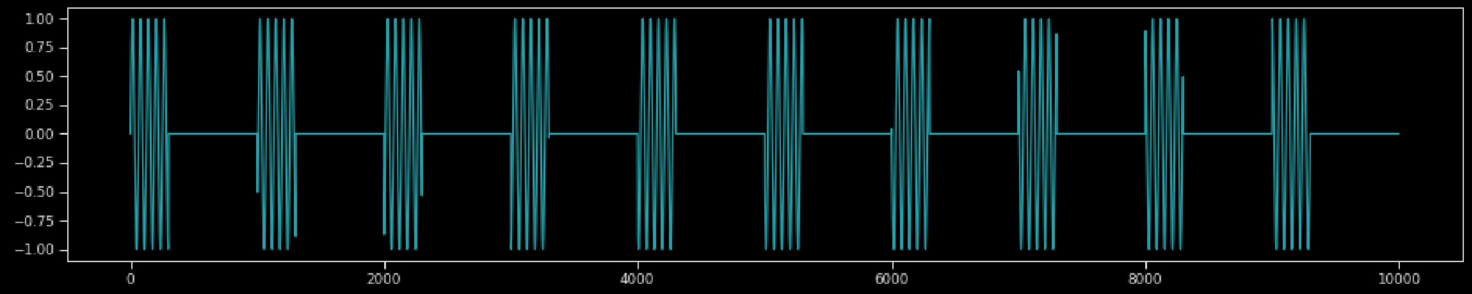
Draw two signals from some common distribution.



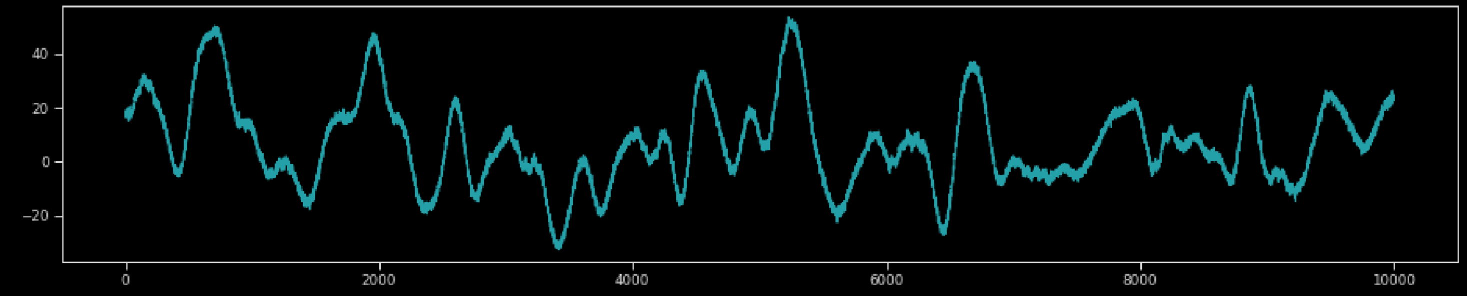
Add in a small but interesting signal to one:



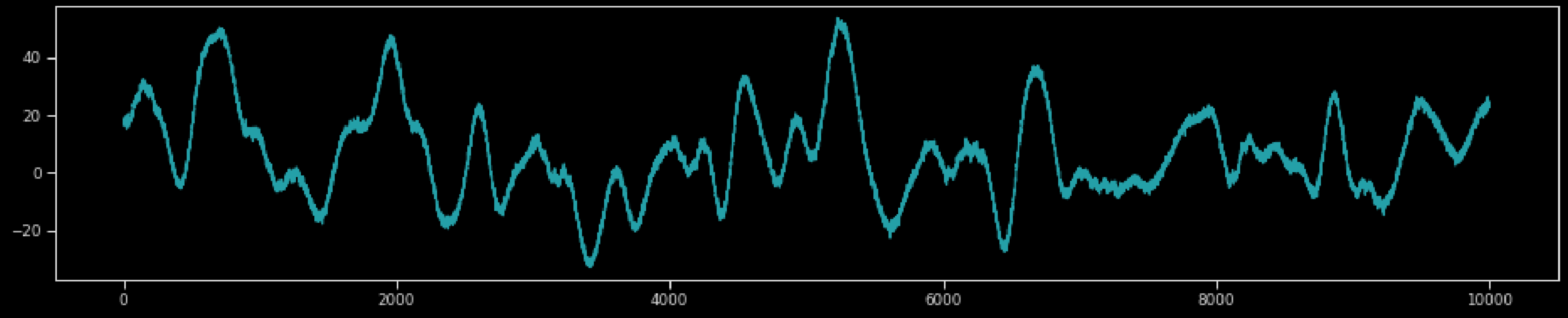
+



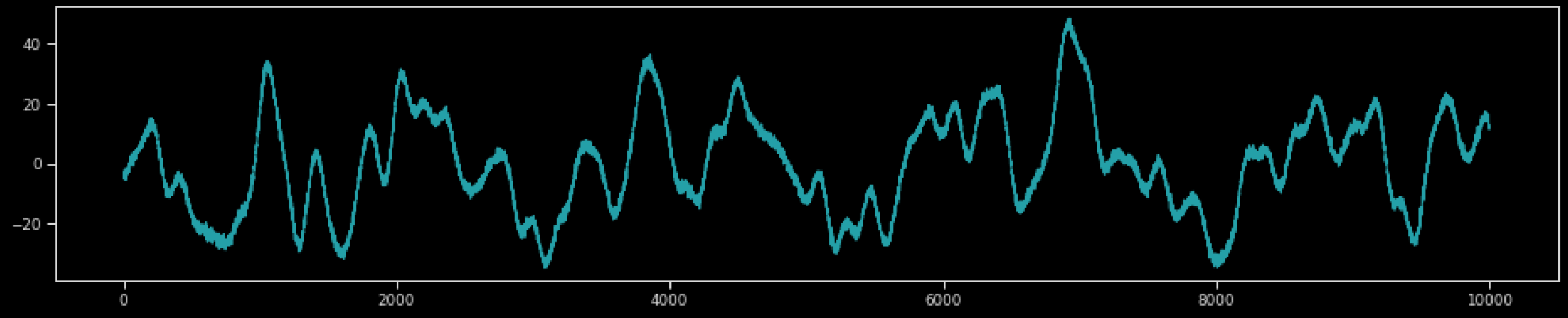
=



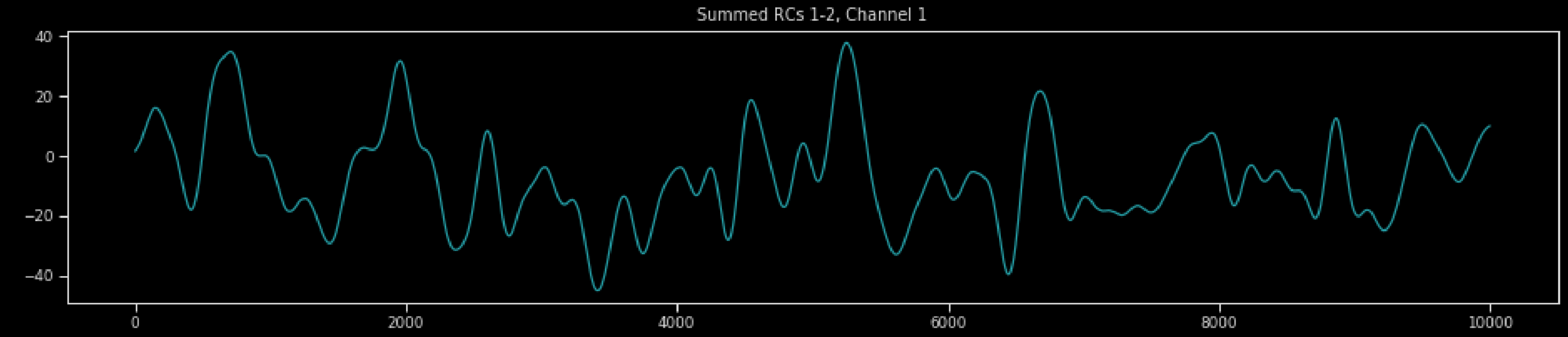
Foreground signal (has special signal):



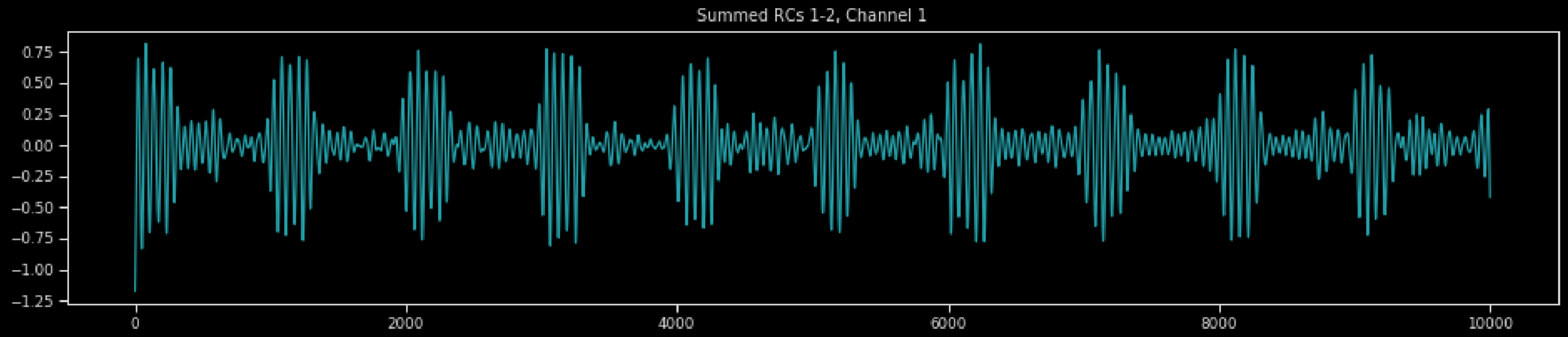
Background signal:



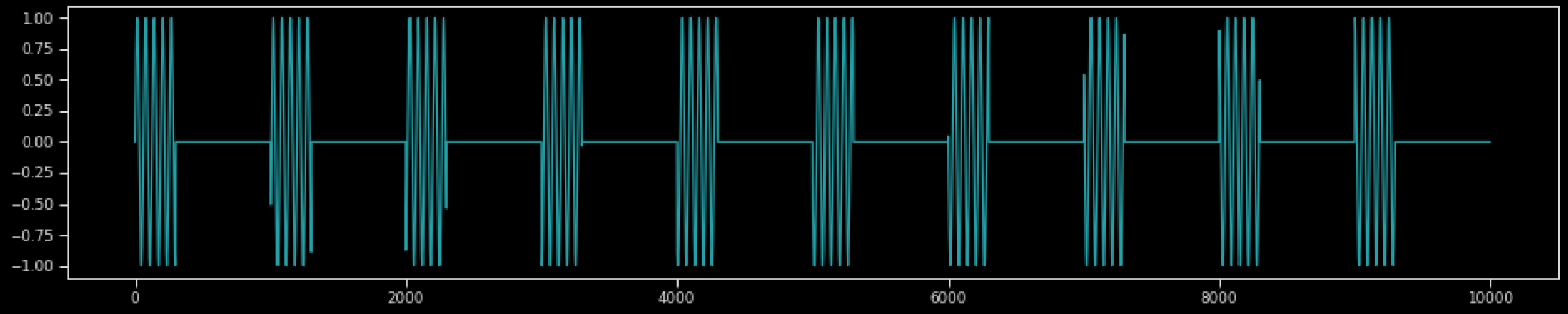
Standard MSSA run on foreground:



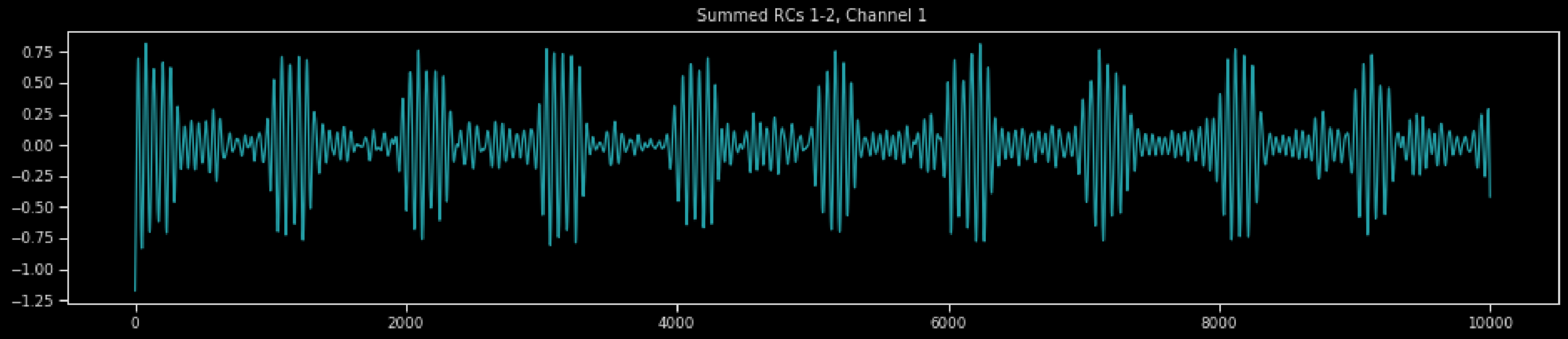
cMSSA run on foreground (w/ background):



Original:



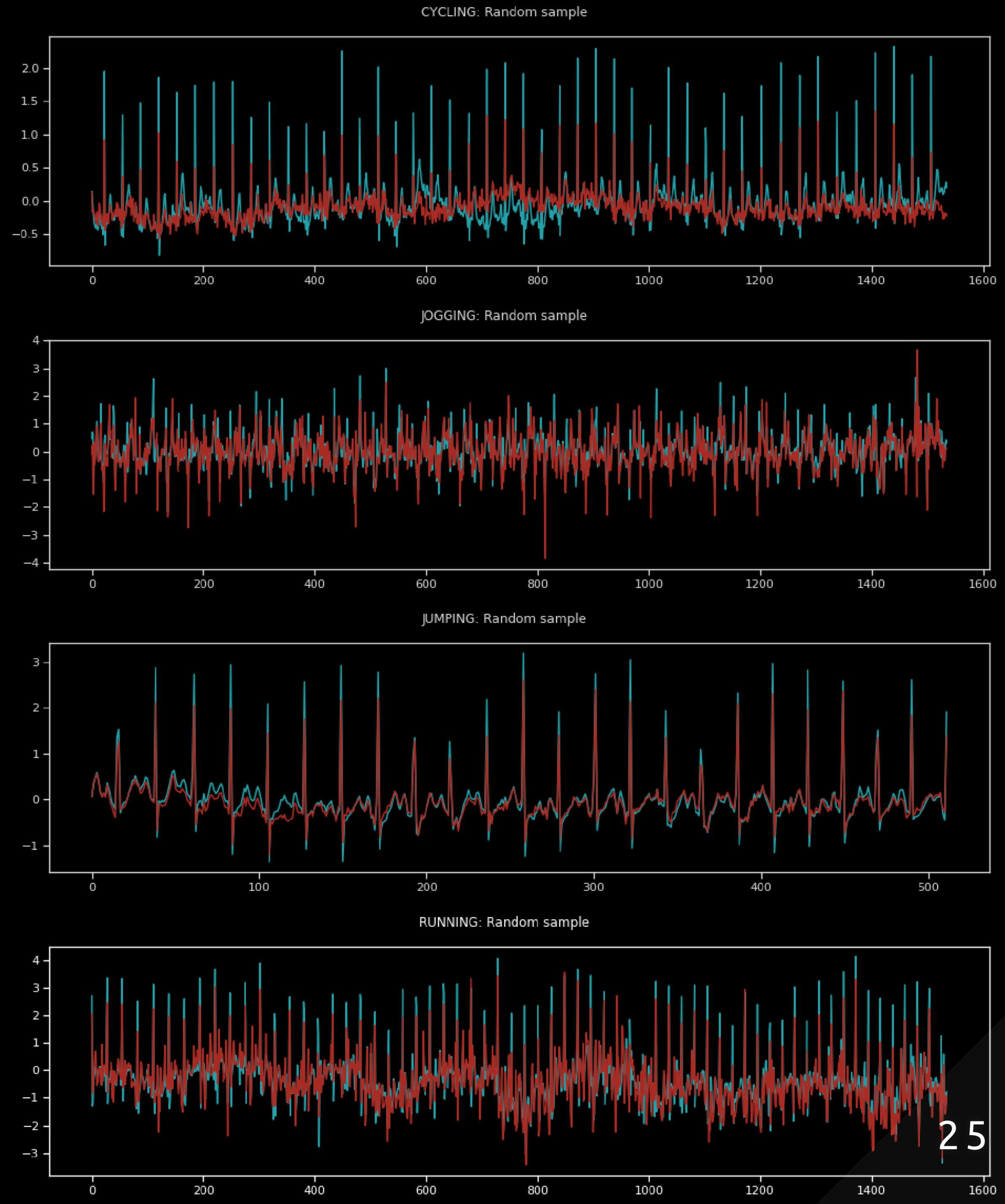
cMSSA run on foreground (w/ background):



Experiments

MHEALTH dataset

Researchers collected 2-lead ECG signals while subjects performed a variety of physical activity.



Task

- Cluster the time series data by activity:
  - Jumping
  - Cycling
  - Jogging
  - Running
- Evaluate against gold truth clusters
  - Metric: BCubed [Amigó 2009].
  - Outputs precision, recall, and F1.

Models

- **Model-free:** Cluster time signals as they are.
- **MSSA:** For various settings of W and K.
- **cMSSA:** MSSA + 4 αs discovered by our procedure.

ECG results

Model	W	K	Space	P / R / F1
Model-free	-	-	-	52.65 / 55.25 / 53.62
MSSA (best)	16	12	PC	54.24/ 60.38 / 57.14
cMSSA (best) ($\alpha = 25.98$)	128	20	PC	**67.12 / 74.63 / 70.67**
MSSA (average)	-	-	-	44.05 / 50.96 / 47.19
cMSSA (average)	-	-	-	58.87 / 65.48 / 61.97

Best cMSSA run uses only **~7.8%** of available basis!

(By comparison, best MSSA run uses **37.5%**).

Results on 8-channel electromyogram (EMG) data.
Goal is to cluster 20 activities.

Model	W	K	Space	P / R / F1
Model-free	-	-	-	25.13 / 27.50 / 26.26
MSSA (best)	8	18	RC	25.51 / 32.50 / 28.58
cMSSA (best) ($\alpha = 3.10$)	16	6	PC	**34.04 / 41.56 / 37.43**
MSSA (average)	-	-	-	21.97 / 23.65 / 22.75
cMSSA (average)	-	-	-	26.65 / 29.62 / 28.02

MSSA: ~28.1% utilization (18 out of 64)

cMSSA: ~4.7% utilization (6 out of 128)

Physical Interpretation

$$\underset{V \in \mathbb{R}^{W \times D}}{\arg\max}\, P(X * V) - \alpha P(Y * V)$$

under the constraint $\|V\|_F = 1.$

$P(\cdot)$ is the power of a univariate signal.

Shifts the goal from finding signals that **explain the most** variance to signals that **matter the most** to the analyst.

cMSSA is available as a Python package at:

https://github.com/aadah/cMSSA

Thank you

Other values for α discovered automatically on synthetic data:

- 1.23
- 4.53
- 20.09
- 50.94