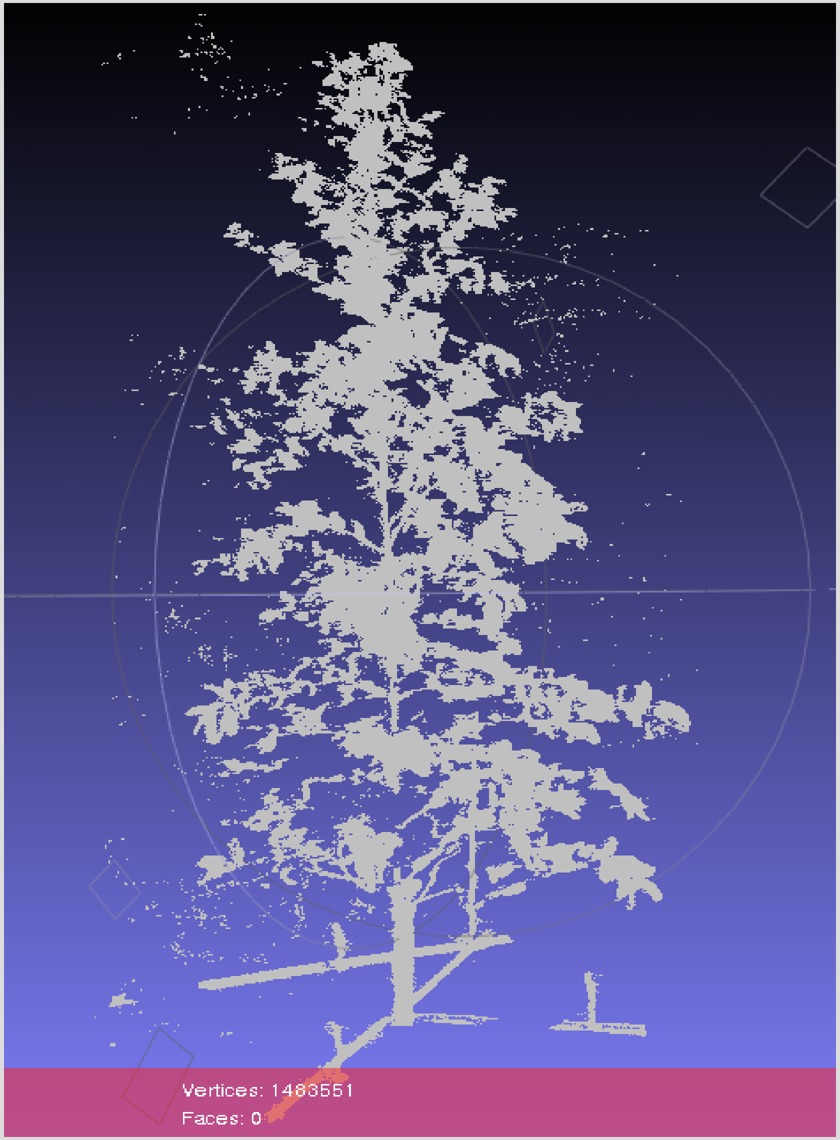
Distinction de feuilles et de branches sur des arbres scannés



Plan de la présentation

- I. Présentation de l'équipe et du sujet
- II. Intuition du spectral clustering
- III. Questions inhérentes au sujet

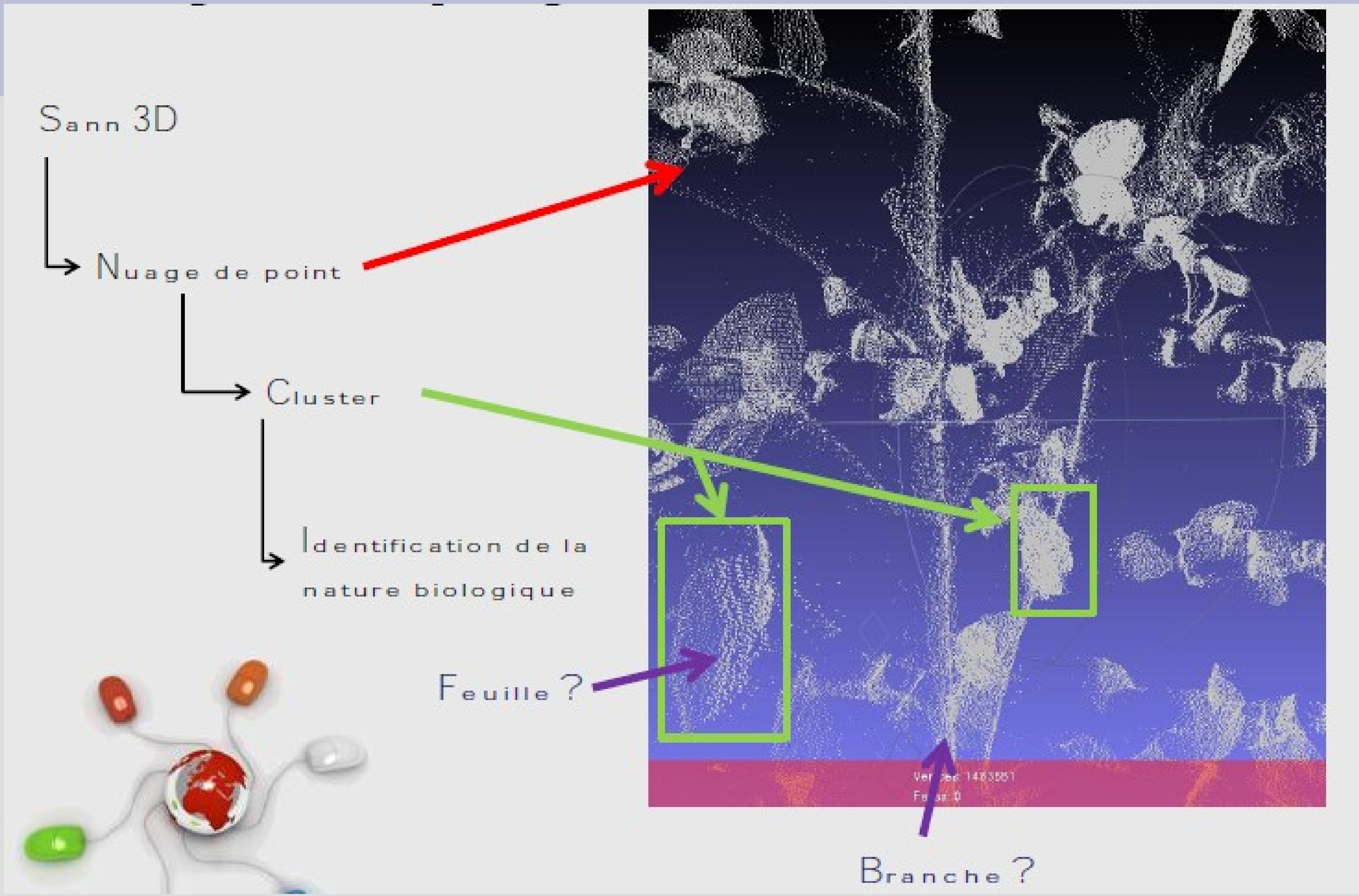
Plan de la présentation

- I. Présentation de l'équipe et du sujet
- II. Intuition du spectral clustering
- III. Questions inhérentes au sujet

I.L'équipe

L'équipe est composée de :
Antoine Gay (IF)
Lionel Guez (MMIS)
Jeremy Kennel (ISI)

I. Le sujet du projet



Plan de la présentation

- I. Présentation de l'équipe et du sujet
- II. Intuition du spectral clustering
- III. Questions inhérentes au sujet

II. Le spectral clustering à travers un exemple

- Matrice Laplacienne

$$L_{i,j} = \begin{cases} -poids \quad de \quad l'arete \quad entre \quad i \quad et \quad j & si \quad i \neq j \\ degre \quad de \quad l'element \quad i & si \quad i = j \end{cases}$$

- Exemple

$$\begin{pmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

II. Le spectral clustering à travers un exemple

- Algorithme du spectral clustering

0. Choisir un critère de similarité

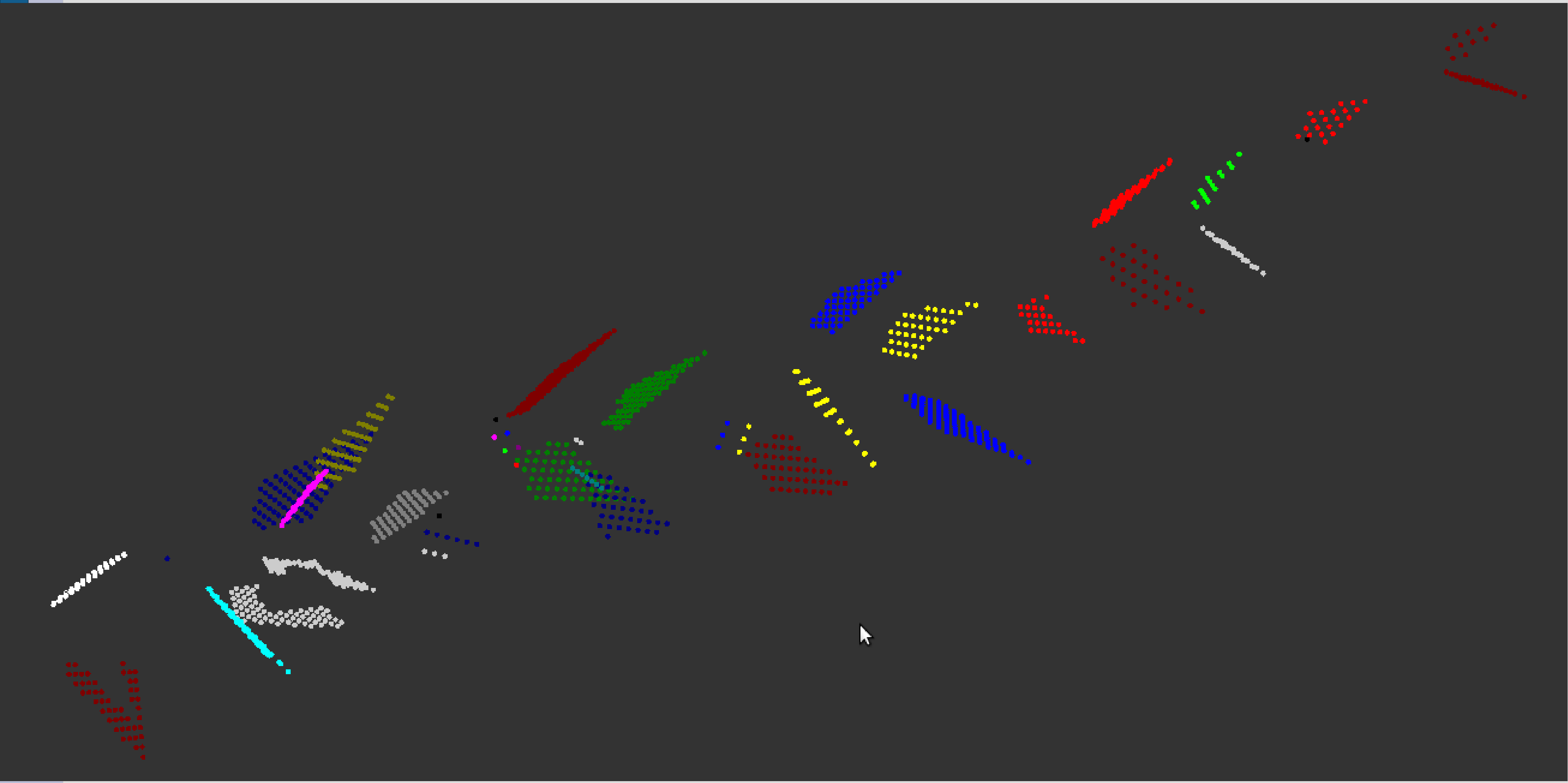
1. Construire la matrice Laplacienne

2. Calculer les k-vecteurs propres, $U_1,\ldots,U_k$, associés au k plus petites valeurs propres.

2bis. Concatener ces k vecteurs propres en une matrice U de taille k ligne et nombre de points colonne.

3. En appliquant l'algorithme des k-means sur les lignes de la matrice U, identifier à quel cluster appartient quel point.

II. L'algorithme fonctionne



Plan de la présentation

- I. Présentation de l'équipe et du sujet
- II. Intuition du spectral clustering
- III. Questions inhérentes au sujet

III. Le graphe – une portion de peuplier

III. Les différents graphes

- Pour créer un graphe à partir d'un nuage de points, il existe différentes stratégies :

- ε-voisinage
- K-voisinage
- K-mutuels voisinages
- Voisinages mixtes.
- Graphes complets(distance de diffusion)

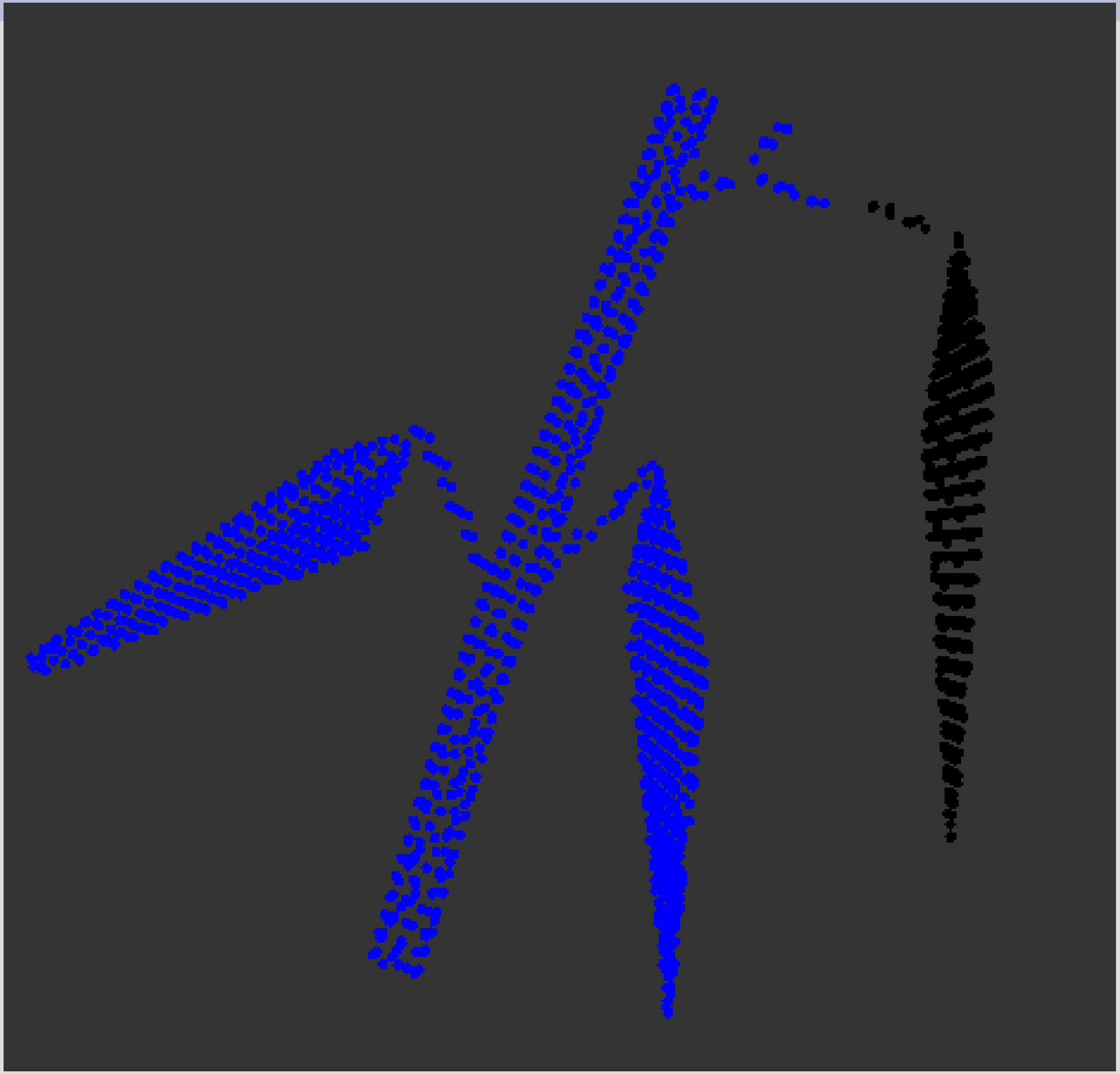
III. ε-voisinage

- Avantage :
  - simple,intuitif

- Défaut :
  - peut isoler des points
  - difficile de choisir un bon ε

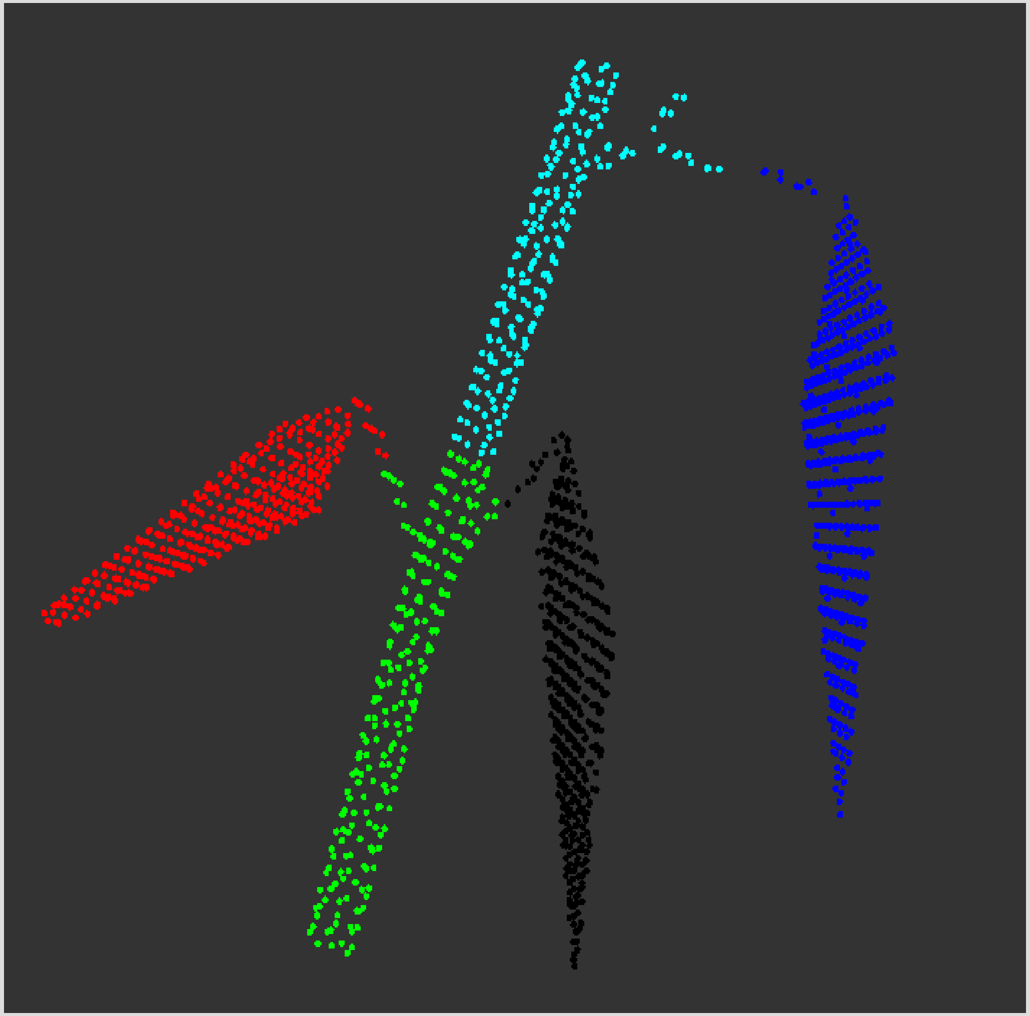
III. ε-voisinage

- Epsilon =0.5
- 2 clusters



III. ε-voisinage

- Epsilon =0.5
- 5 clusters

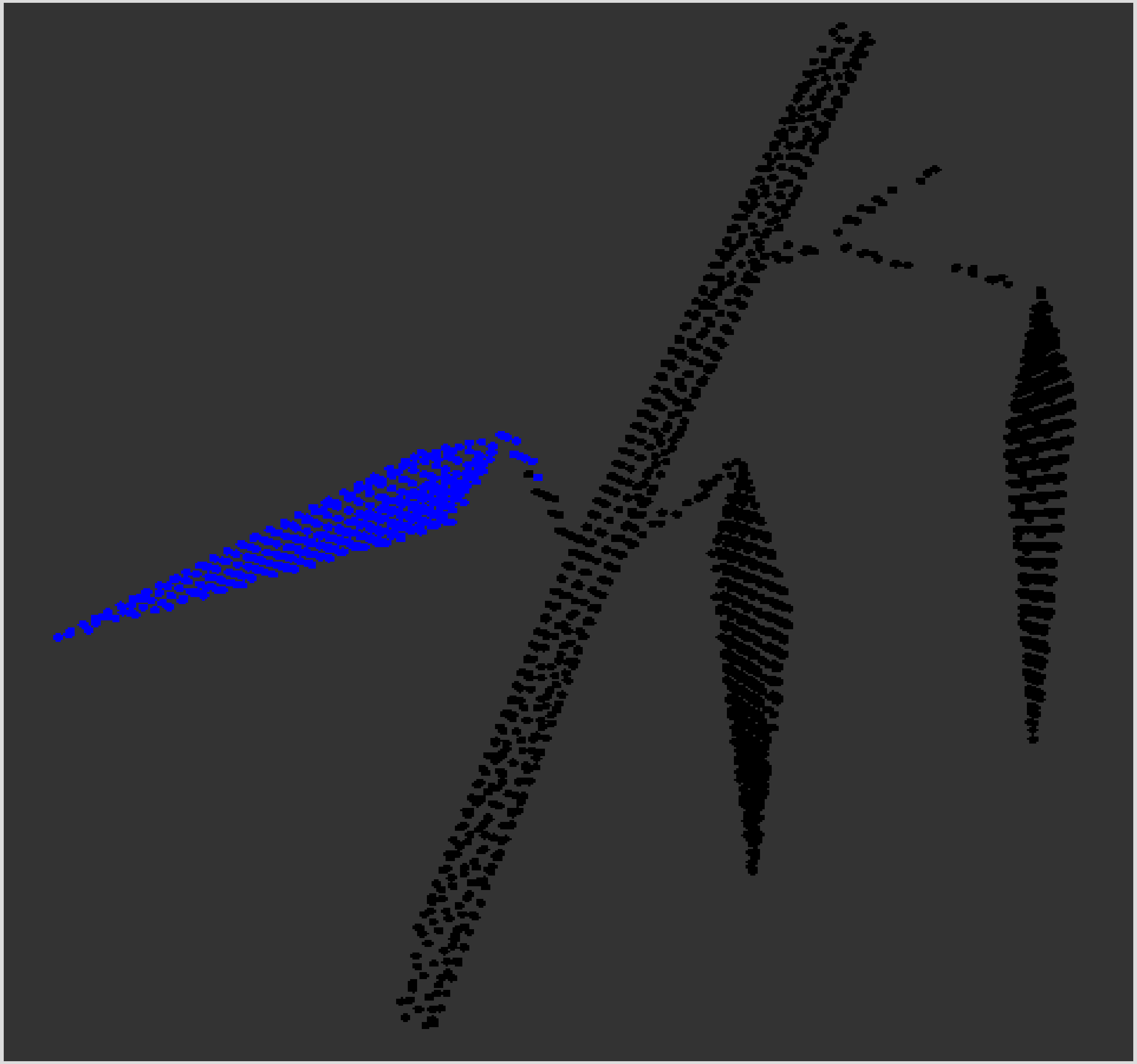


III. K-voisinage

- Avantages :
  - jamais de points isolés
  - Indépendant du pas du laser => pas de difficulté pour trouver un bon k
- Défauts :
  - sépare mal le graphe en composantes connexes
  - Relativement long à éxécuter

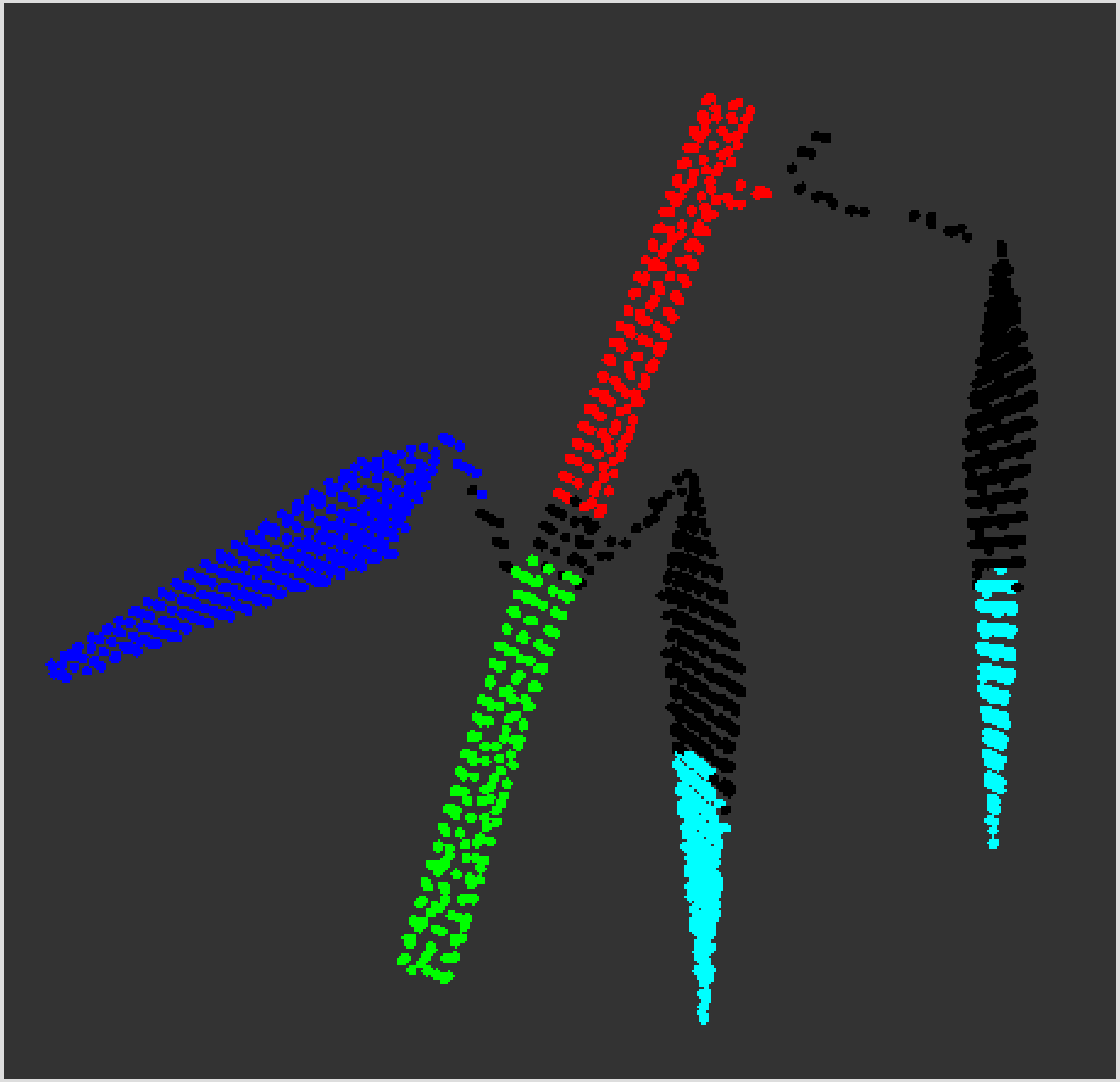
III. K-voisinage

- K = 4
- 2 clusters



III. K-voisinage

- K = 4
- 5 clusters

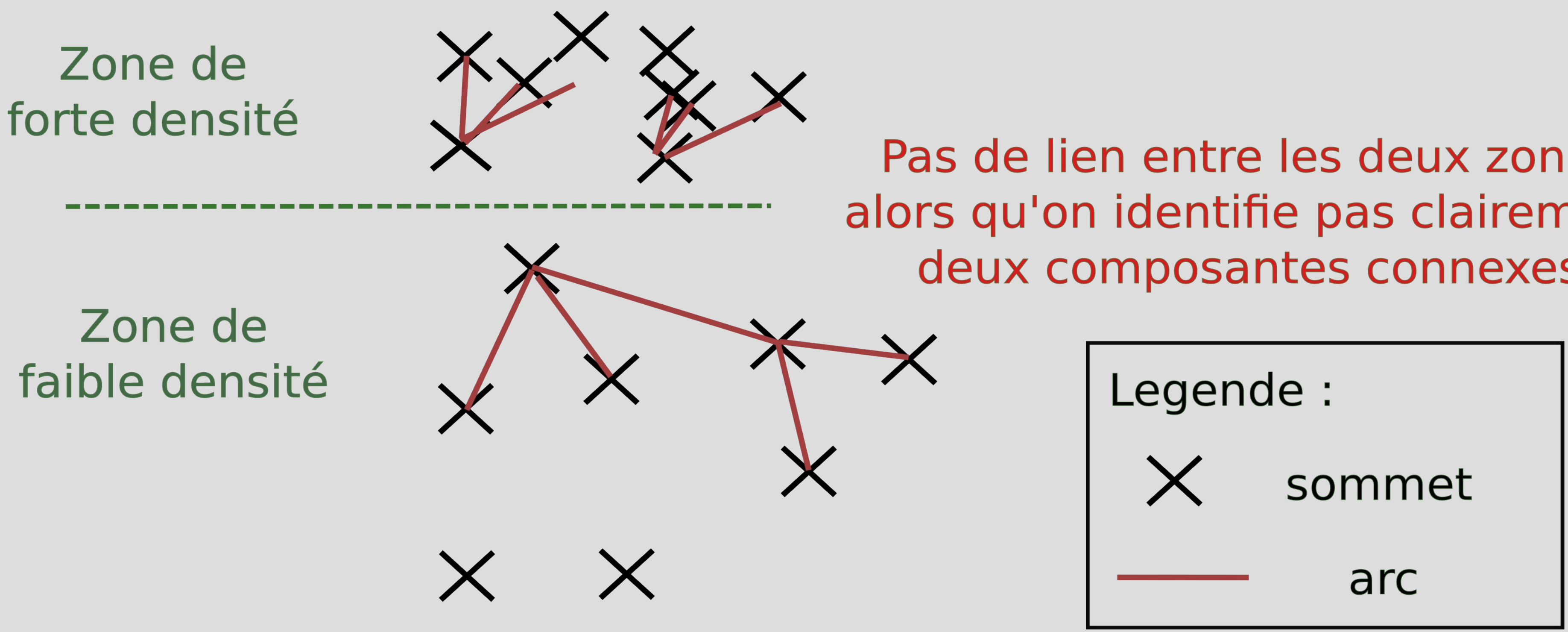


III. K-mutuel voisinage

- Avantage :
  - Sépare bien le graphe en composantes connexes sans trop de points isolés
  - Ne connecte pas les zones de densité différentes
- Défauts :
  - Relativement lent à executer

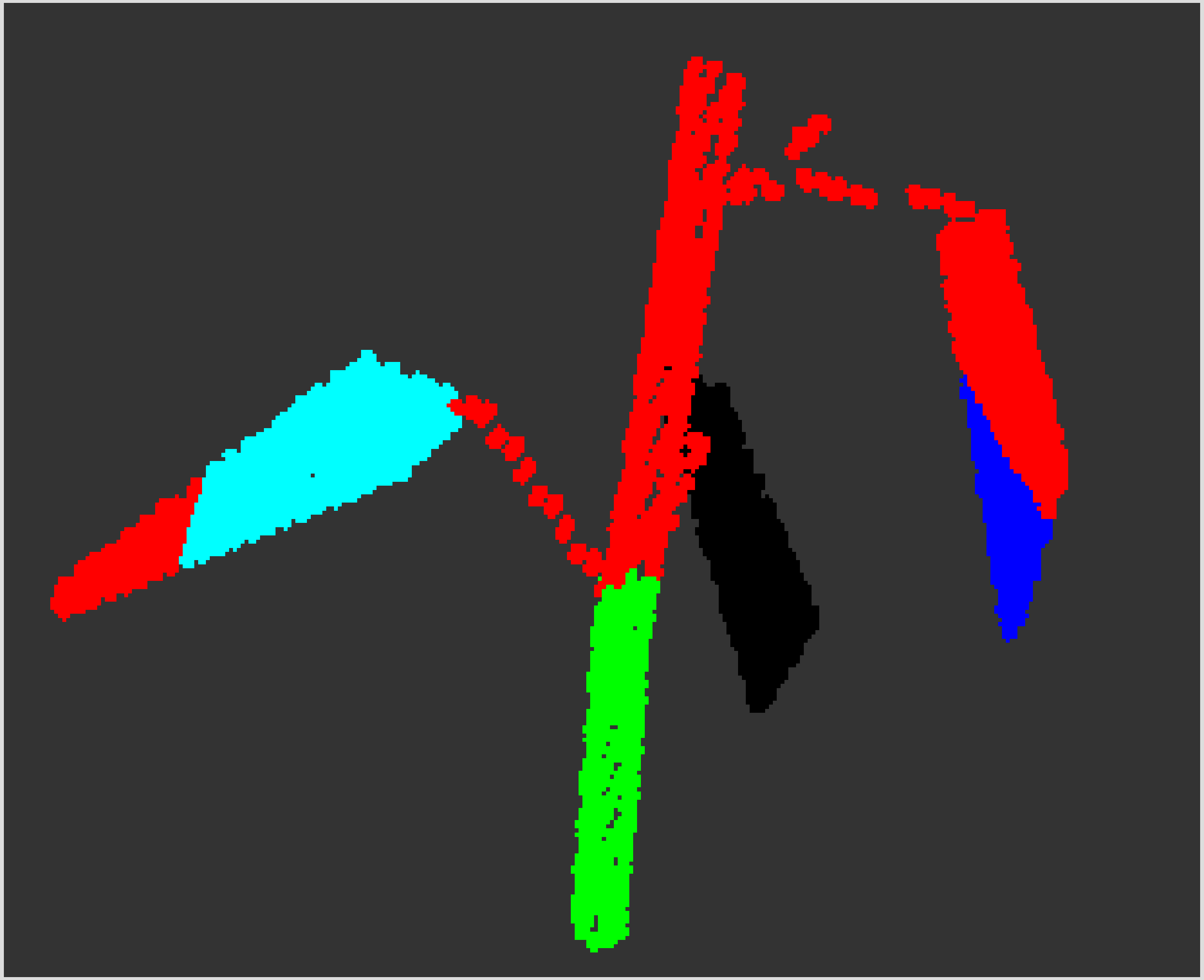
III. K-mutuel voisinage

Représentation partielle du graphe
pour le k mutuel voisinage avec k=3



III. K-mutuel-voisinage

- K = 4
- 5 clusters



III. Voisinage mixte

- Avantages :
  - matrice bien creuse
  - Sépare bien en composante connexes
- Défauts
  - Points isolés
  - Difficultés pour trouver un bon ε

III. Distance de diffusion

Distance euclidienne

Formule : Pour M de coordonnées (x_M, y_M, z_M) et N de coordonnées (x_N, y_N, z_N),

on a $dist_{eucli}(M,N) = \sqrt{(x_M - x_N)^2 + (y_M - y_N)^2 + (z_M - z_N)^2}$

Une distance de diffusion

Formule : Pour M et N deux points de l'espace,

on a $dist_{diffu}(M,N) = exp(-\frac{(dist_{eucli}(M,N))^2}{\epsilon})$

III. Distance de diffusion

- Matrice epsilon diffusion

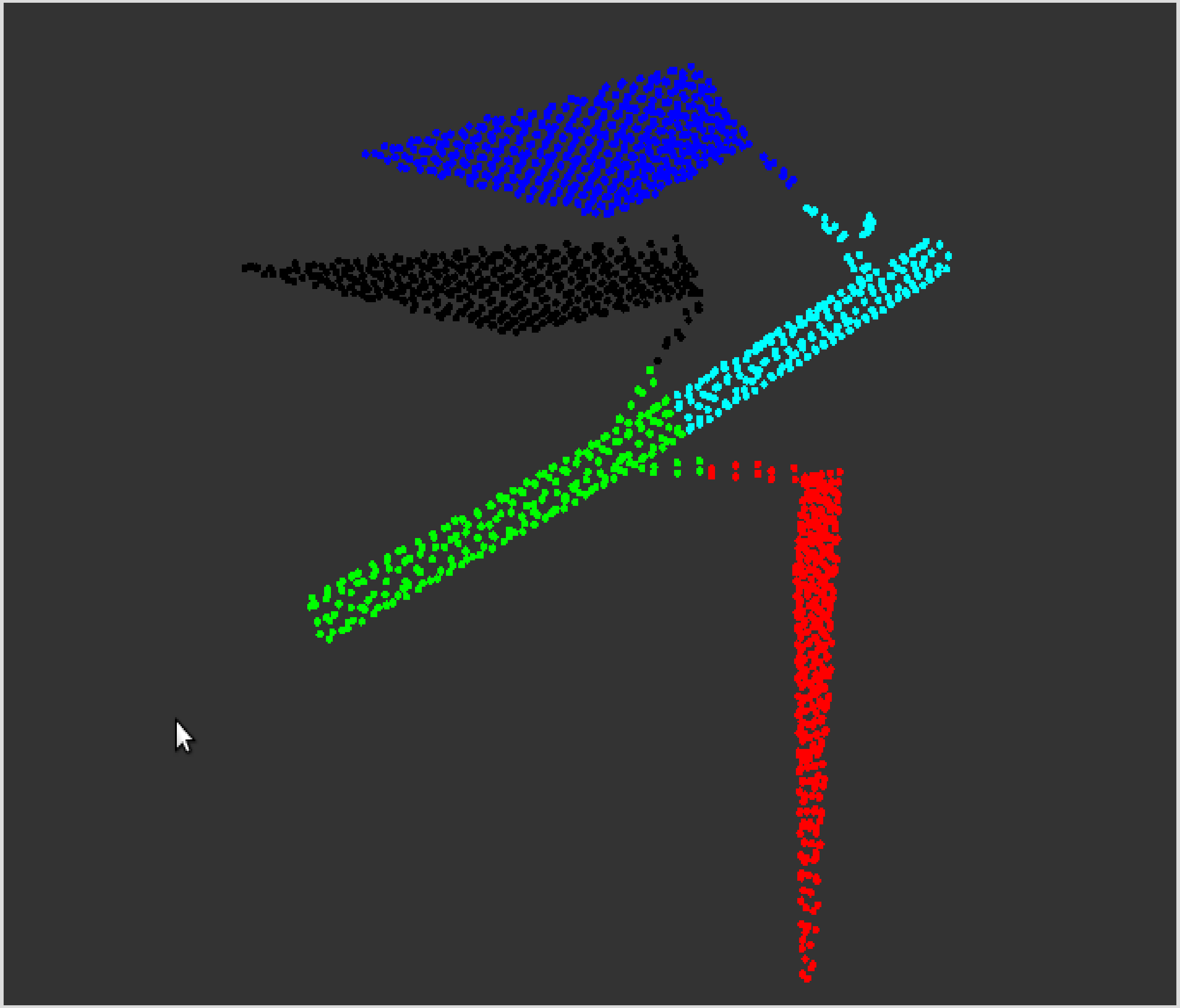
```
1, -0.333333, -0.333333, -0.333333, 0, 0, 0, 0,
-0.5, 1, -0.5, 0, 0, 0, 0, 0,
-0.333333, -0.333333, 1, -0.333333, 0, 0, 0, 0,
-0.5, 0, -0.5, 1, 0, 0, 0, 0,
0, 0, 0, 0, 1, -0.333333, -0.333333, -0.333333,
0, 0, 0, 0, -0.333333, 1, -0.333333, -0.333333,
0, 0, 0, 0, -0.333333, -0.333333, 1, -0.333333,
0, 0, 0, 0, -0.333333, -0.333333, -0.333333, 1,
```

- Matrice de la distance de diffusion

```
1, -0.908241, -0.0917586, -1.13238e-10, -0, -0, -0, -0,
-0.999914, 1, -8.63752e-05, -2.37687e-16, -0, -0, -0, -0,
-0.0216503, -1.85117e-05, 1, -0.978331, -0, -0, -0, -0,
-2.73102e-11, -5.20686e-17, -1, 1, -0, -0, -0, -0,
-0, -0, -0, -0, 1, -0.746501, -0.2065, -0.0469989,
-0, -0, -0, -0, -0.211739, 1, -0.511023, -0.277239,
-0, -0, -0, -0, -0.052504, -0.458081, 1, -0.489415,
-0, -0, -0, -0, -0.0159356, -0.331408, -0.652656, 1,
```

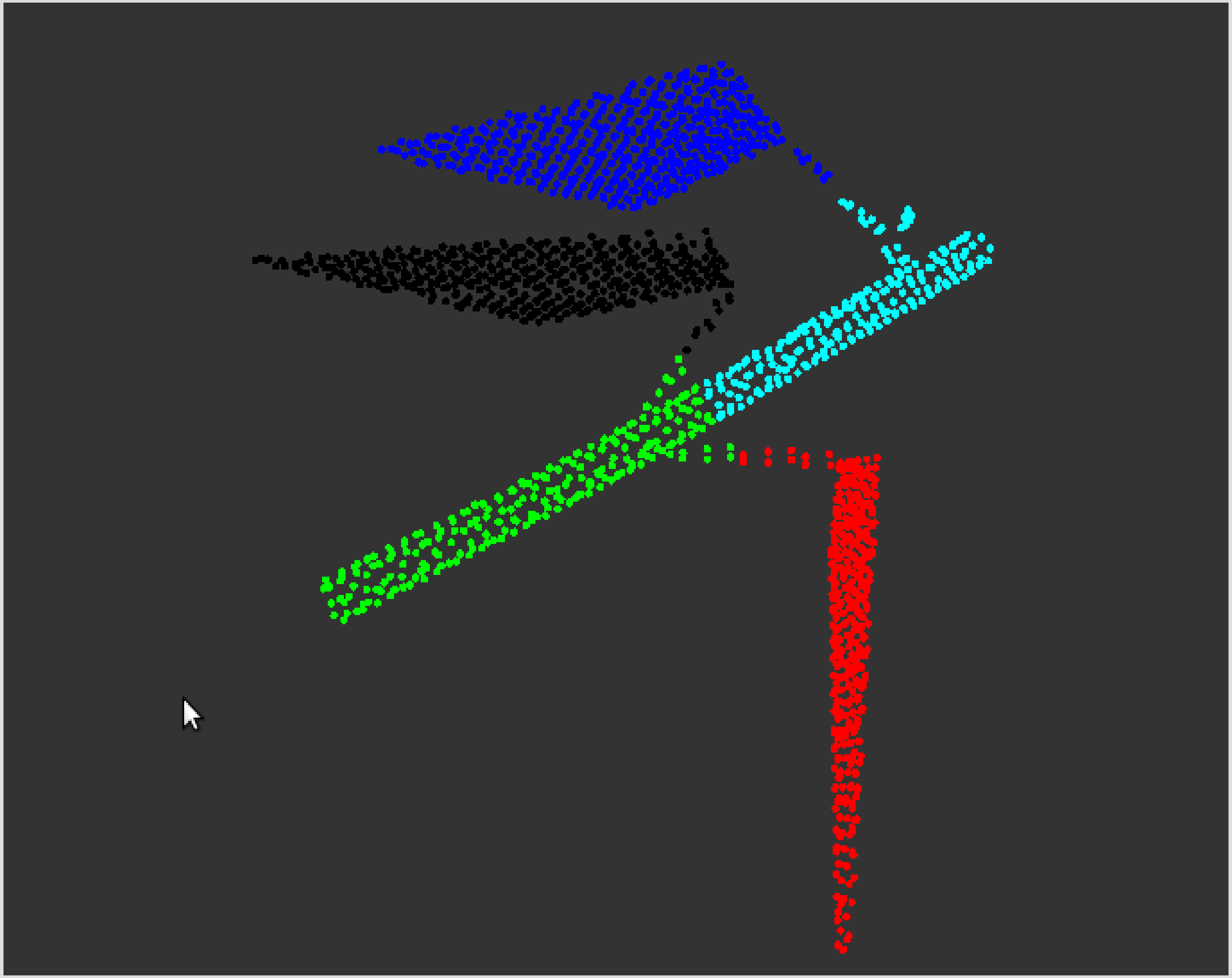
III. Distance de diffusion

- 5 clusters



III. Distance de diffusion

- 4 clusters



III. Le nombre de clusters

III. Taille des fichiers

- Dépassement de capacité

Conclusion

Nous conservons :

- l'epsilon voisinage
- La distance de diffusion