

# Directional-iDBSCAN

*first look at LIME data*

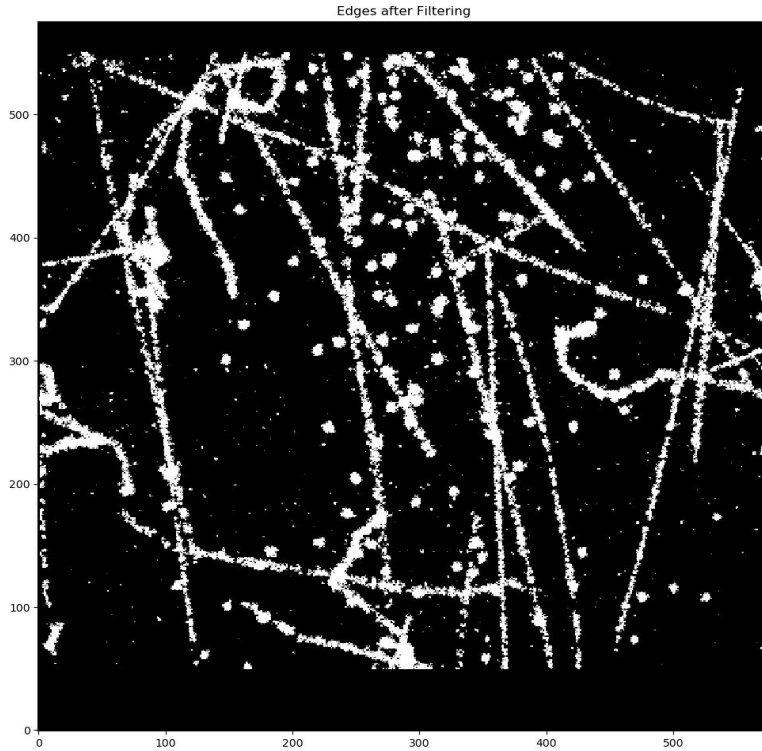
**Igor Pains**

Igor Abritta and Rafael A Nobrega

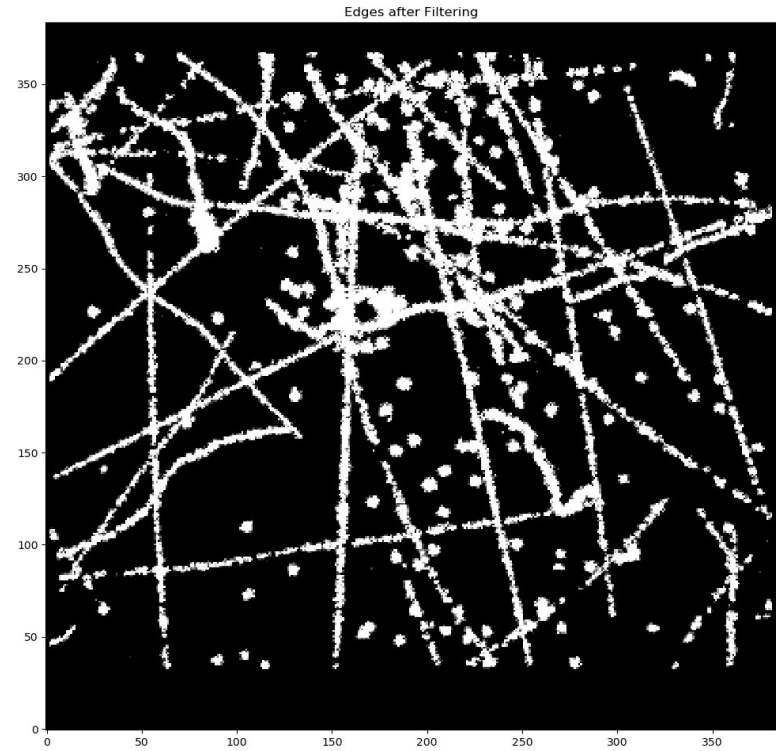
# Última apresentação

- Novo banco de dados para análise com o iDDBSCAN.
  - Runs: 4433 (cosmic); 4455 (Fe55) - 200 eventos cada.
  
- Mudança de branch para versão mais atual do algoritmo (*autumn21*).
  - Pequenas mudanças em vários arquivos do reconstruction como correções de bugs e novas funções (selecionar área da imagem a ser clusterizada, vetor customizado de eventos para rodar o `debug_mode = 0, ...`)
  - Mudanças em alguns parâmetros do iDDBSCAN e `rebin = 6`.

# Ocupação no banco de dados



Run 4455 - Event 6 (rebin 4)



Run 4455 - Event 10 (rebin 6)

# Mudanças iDDBSCAN

- Modificação do RANSAC utilizado no *DBSCAN seeding* para a versão com teste com clusters rotacionados. (discutido na última reunião)
- *Isolation seeding* movido para fora do loop. (estava levando muito tempo pra ser feito nesse novo banco de dados)

# Reco files iniciais

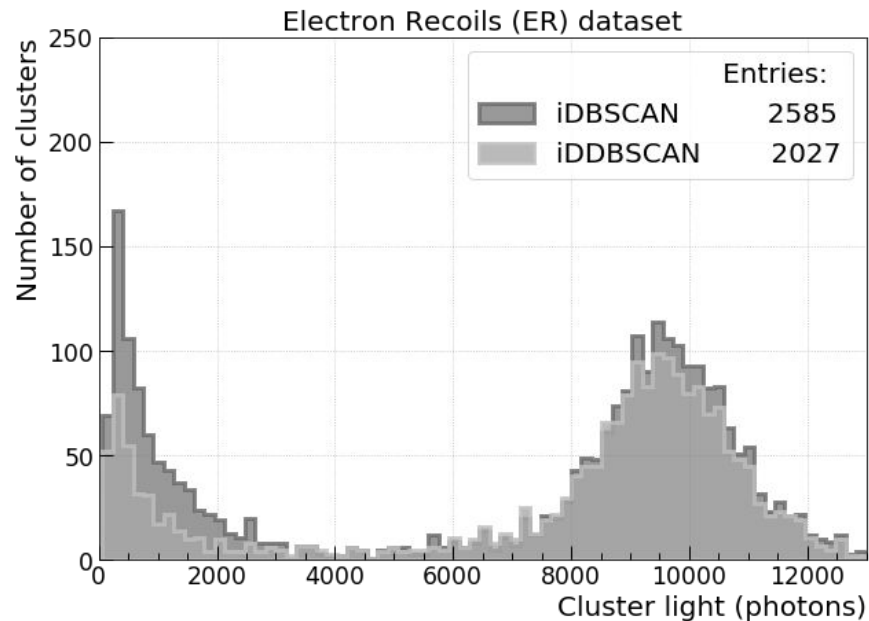
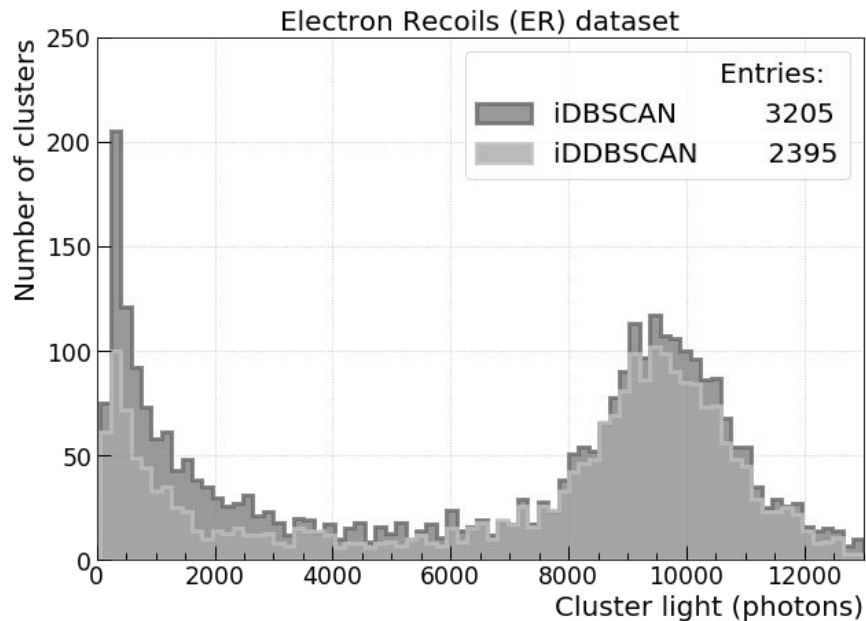
23 lines (21 sloc) | 853 Bytes

```
1 {
2
3 ## DBSCAN seeding
4 'dim'          : '3D',
5 'dbscan_eps'   : 1.1,
6 'dbscan_minsamples' : 5, # this is for 2D #
7
8 ## directional clustering
9 'dir_radius'   : 7.5,
10 'dir_min_accuracy' : 0.8, # minimum accuracy of the RANSAC to save one point of the cluster for the directional search
11 'dir_minsamples' : 20, ## N.B. this is always 2D, and it may differ from dbscan_minsamples (that can be Npix * intensity for 3D)
12 'dir_thickness' : 3.5,
13 'time_threshold' : 300, # seconds
14 'max_attempts'  : 13,
15 'isolation_radius' : 10,
16 'metric'        : 'cityblock', # this is fundamental not to be "euclidean", in order not to merge close and long tracks
17 'metric_params' : None,
18 'algorithm'     : 'auto',
19 'leaf_size'     : 30,
20 'p'             : None,
21 'n_jobs'        : None,
22 'expand_noncore' : True,
23 }
```

- Maiores modificações: eps, min\_samples, dir\_radius, dir\_minsamples, **metric**. (algumas modificações por causa da mudança no rebin)

Run 4433

# Run 4455



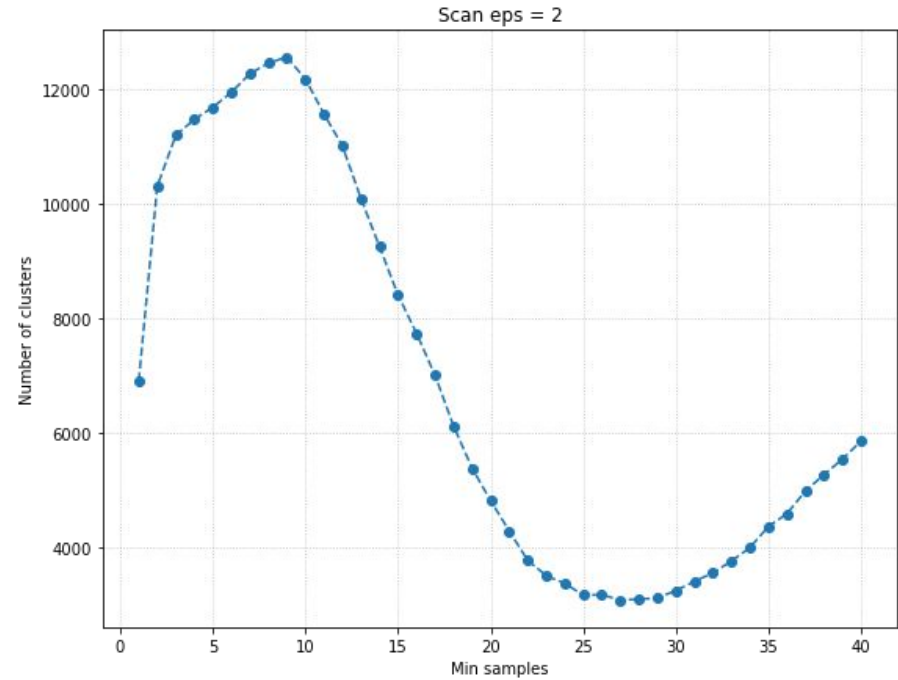
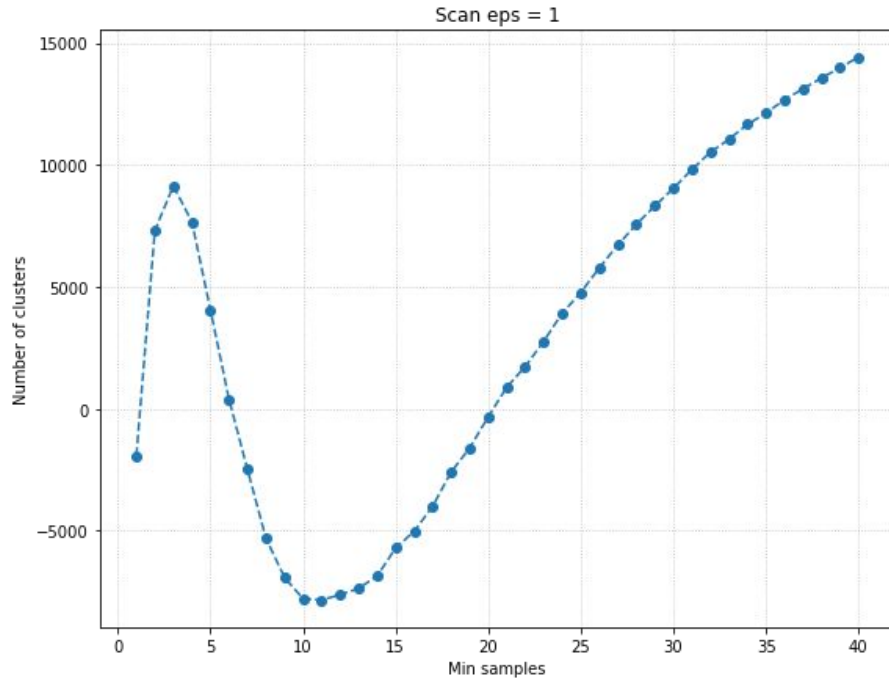
Distribuições de energia no run 4455 sem (esquerda) e com (direita) seleção por slimness ( $> 0.6$ )

# Parâmetros iDBSCAN

- Scan dos parâmetros inspirado no artigo do iDBSCAN (máximo entre clusters encontrados em ER - NRAD)
- Como a métrica “cityblock” (Manhattan distance) é o módulo da diferença das coordenadas de dois pontos, eps deve ser inteiro.
- Scan foi feito nos seguintes intervalos:
  - eps: [1, 2, 3, 4]
  - min\_samples: [1, 2, ..., 40]

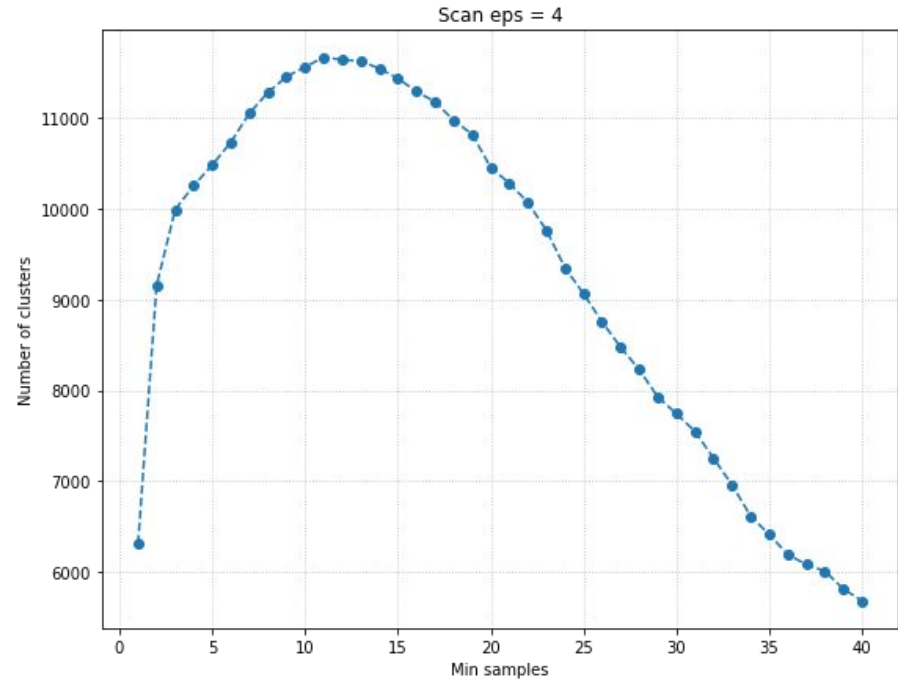
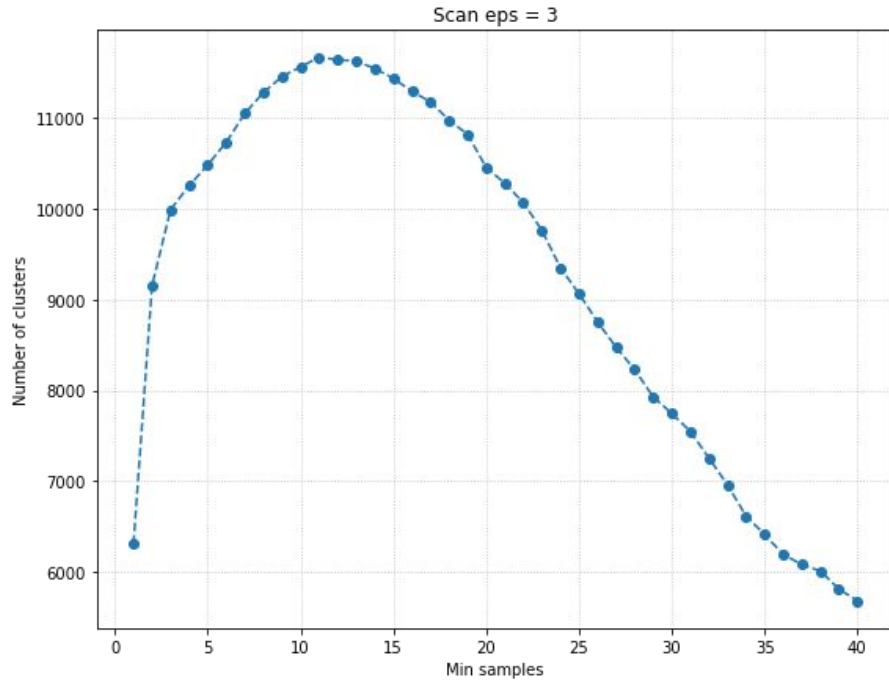


# Parâmetros iDBSCAN



ER - NRAD (Fe55 spots)

# Parâmetros iDBSCAN



ER - NRAD (Fe55 spots)

# Parâmetros iDBSCAN

- Resultados podem (muito provavelmente) estar enviesados por consequência do banco de dados menor.
- Mesma análise ainda deve ser realizada com rebin 4.
- Por enquanto utilizarei  $[\text{eps}, \text{min\_pts}] = [1, 4 \text{ (ou 5)}]$  ou  $[2, 9]$ .

# Próximos passos

- Rodar o algoritmo novamente com os novos parâmetros.
- Refazer análise dos parâmetros com rebin 4 e com métrica *euclidian*. (essa vai demorar)
- Verificar novas condições a serem utilizadas na seleção de clusters dos reco\_files.