多高斯拟合的原理

$$f(x,y,v) = Aexp(-(a(x-x_0)^2 + 2b(x-x_0)(y-y_0) + c(y-y_0)^2 + d(v-v_0)^2))$$

$$a = \frac{cos^2\theta}{2\sigma_x^2} + \frac{sin^2\theta}{2\sigma_y^2}$$

$$b = -\frac{sin2\theta}{4\sigma_x^2} + \frac{sin2\theta}{4\sigma_y^2}$$

$$c = \frac{sin^2\theta}{2\sigma_x^2} + \frac{cos^2\theta}{2\sigma_y^2}$$

$$d = \frac{1}{2\sigma_v^2}$$

A: 幅值

θ: 银经银纬面上的旋转角

(x_0, y_0, v_0):质心

$(\sigma_x, \sigma_y, \sigma_v)$: 三个轴方向上的标准差

采用最小二乘法进行求解

Minimize $F(\vartheta) = \sum_i \rho((f_i(x,y,v) - I)^2)\ i = 1,2,\dots,m$

subject to $lb \le \vartheta \le lb$

scipy.optimize.least_squares

单个3维高斯模型需要拟合8个参数: $\vartheta = (A, x_0, y_0, \sigma_x, \sigma_y, \theta, v_0, \sigma_v)$

拟合过程

初始猜想值及参数边界的确定：

- 参数初始值：采用LDC检测得到的云核的形态参数确定

  $(A, x_0, y_0, \sigma_x, \sigma_y, \theta, v_0, \sigma_v)$=(Peak, Cen1, Cen2, Size1, Size2, 0, Cen3, Size3)

- 参数边界值：

  peak_range = 3　　# 峰值范围[A -3, A +3]

  peak_low = 5 * 0.23　　# 拟合A 的最低值

  sigma_t = 1　　# 质心范围：[x_0 - sigma_t * σ_x, x_0 + sigma_t * σ_x]

  s_time = 1.5　　# σ_x 偏移的倍数：[σ_x * (s_time - 1), σ_x * s_time]

  s_low = 1　　# σ_x的最低值

拟合过程

相互重叠云核的确定：

$$d_{cen} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

$$d_{size} = \sigma_i + \sigma_j$$

$$\sigma_i = \sqrt{{\sigma_{xi}}^2 + {\sigma_{yi}}^2}$$

→相互不重叠

d_{cen} > multi * d_{size} or $|v_i - v_j|$ > multi $*(\sigma_{vi} + \sigma_{vj})$

→相互重叠

$d_{cen} \leq$ multi * d_{size} or $|v_i - v_j| \leq$ multi $*(\sigma_{vi} + \sigma_{vj})$

确定哪些云核相互重叠：

- 计算每两个云核是否重叠，得到邻接矩阵
  第i个核与第j个核相互重叠，touch_mat[i, j]=1

- 根据邻接矩阵，同一棵树上面的即为相互重叠的云核
  from scipy.sparse.csgraph import connected_components

newarr = csr_matrix(touch_mat)
n_components, labels = connected_components(newarr)

```
graph = [[ 0, 1, 1, 0, 0],
         [ 0, 0, 1, 0,0 ],
         [ 0, 0, 0, 0, 0],
         [0, 0 , 0, 0, 1],
         [0, 0, 0, 0, 0]]
```

```
graph = csr_matrix(graph)
print(graph)
    (0, 1)    1
    (0, 2)    1
    (1, 2)    1
    (3, 4)    1
```

n_components, labels =
connected_components(csgraph=graph, directed=False, return_labels=True)
n_components = 2
Labels = array([0, 0, 0, 1, 1], dtype=int32)

拟合过程

- 首先根据重叠云核的个数动态生成对应的多高斯模型;
- 根据LDC检测核表确定对应的拟合参数初始值及参数边界;
- 对拟合结果整理;
- 最终的拟合核表如下:

[ID, Peak1, Peak2, Peak3, Cen1, Cen2, Cen3, Size1, Size2, Size3, theta , Peak, Sum, success, cost]

ID: 分子云核编号

Peak1,Peak2, [Peak3]: 峰值位置坐标

Cen1, Cen2, [Cen3]: 质心位置坐标

Size1, Size2, [Size3]: 云核的轴长[FWHM=2.3548*sigma]

theta: 云核在银经银纬面上的旋转角

Peak: 云核的峰值强度

Sum: 云核总流量(按照表达式积分)

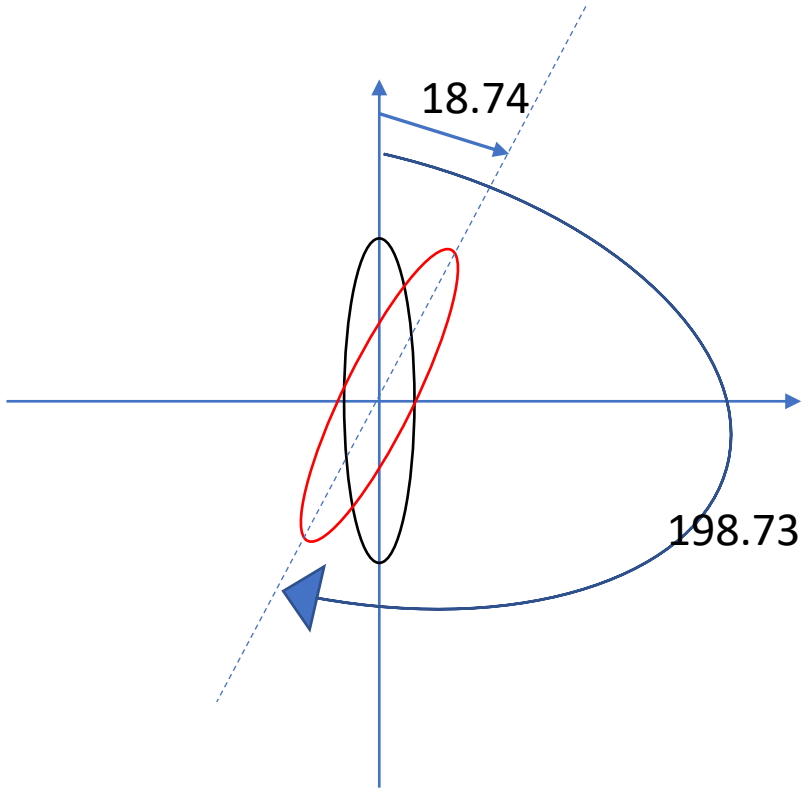
success: 拟合是否成功

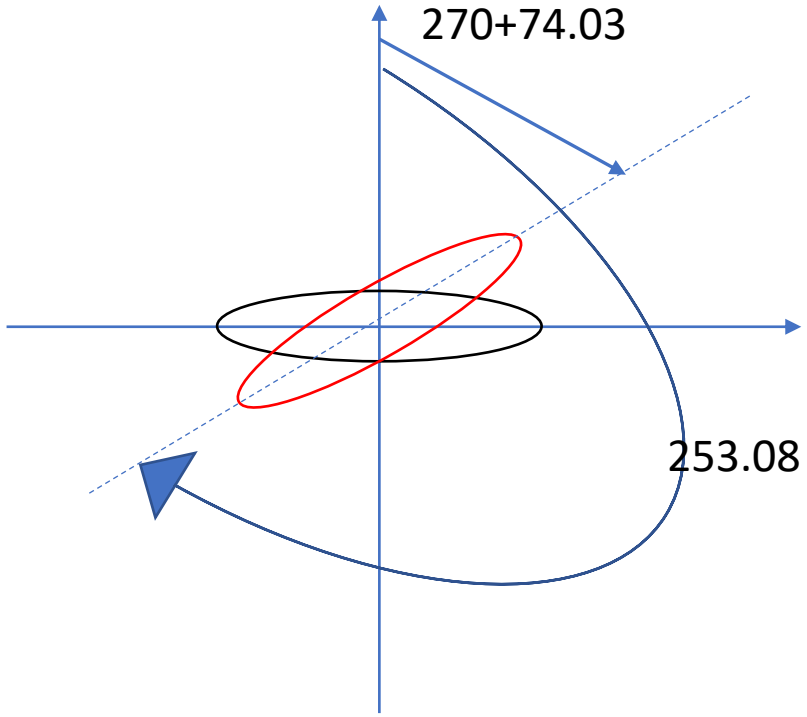
cost: 拟合函数和实际数据间的误差

- 当size_x>size_y
- Mod(s_angle,180) = mod(fit_angle,180)

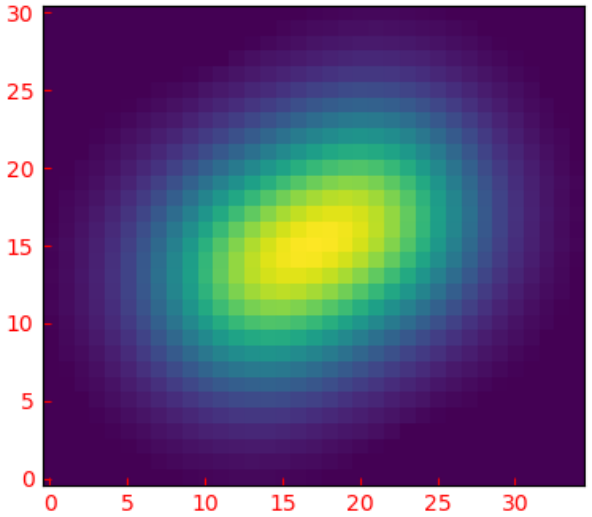
- 当size_x<size_y
- Mod(s_angle,180) = mod(fit_angle+90,180)

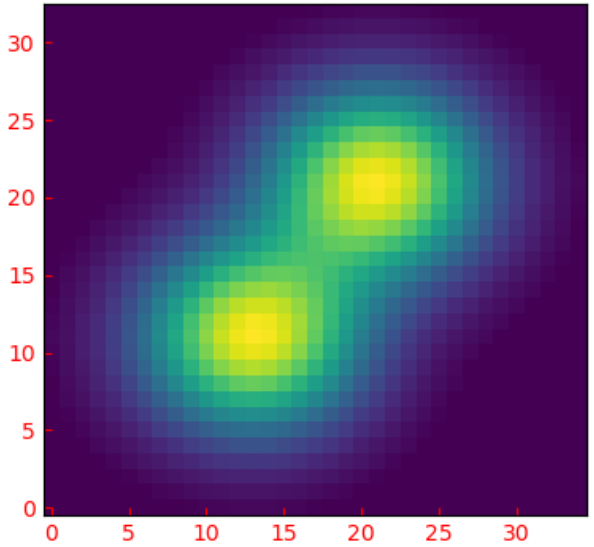
s_angle	fit_angle	size_max	size_min	size_x	size_y
198.73	18.74	5.32	3.05	5.46	3.28
253.08	343.06	4.77	3.33	3.54	4.92

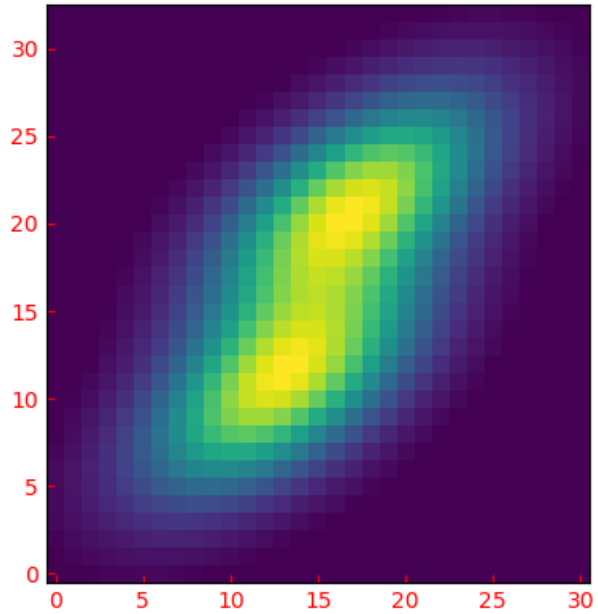




多高斯成分拟合在仿真分子云数据上的实验



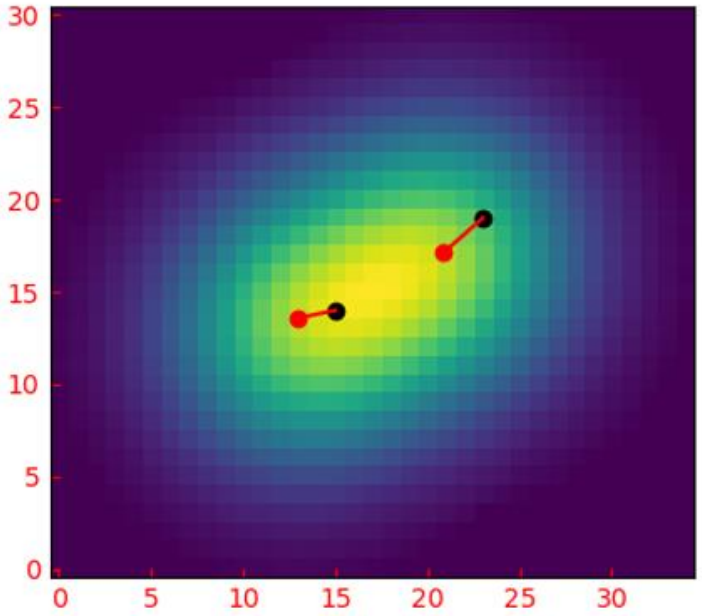


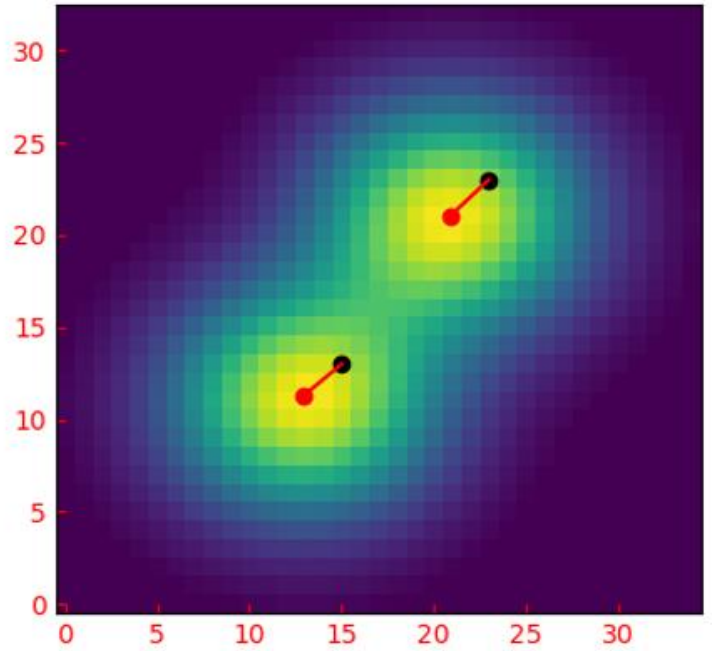


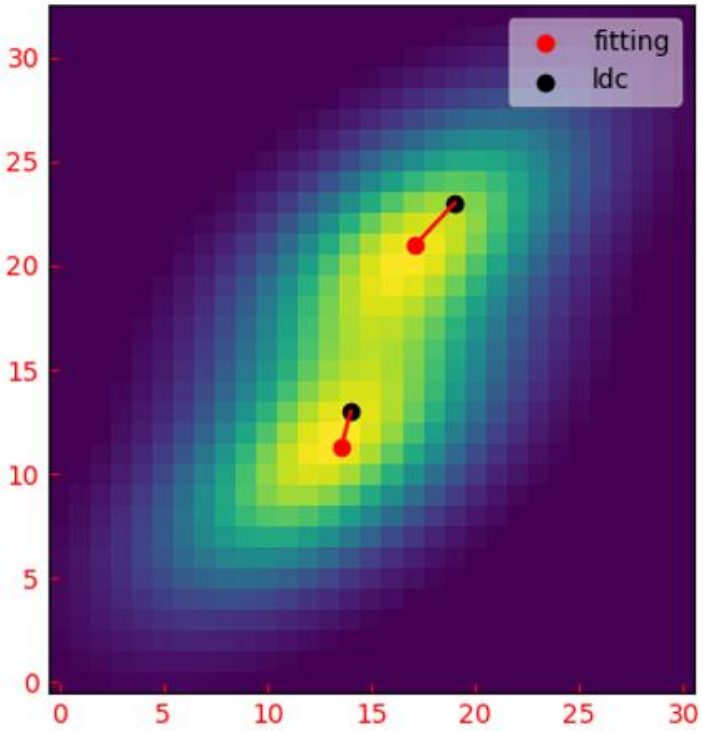
ID	Peak1	Peak2	Peak3	Cen1	Cen2	Cen3	Size1	Size2	Size3	theta	Peak	Sum	Volume
0.00000	51.00000	51.00000	51.00000	51.00000	51.00000	51.00000	14.12900	7.06400	11.77400	36.00000	3.80000	5386.36700	15455.00000
1.00000	41.00000	47.00000	43.00000	41.00000	47.00000	43.00000	14.12900	7.06400	11.77400	36.00000	3.80000	5386.36700	15455.00000

仿真云核的参数

多高斯成分拟合在仿真分子云数据上的实验







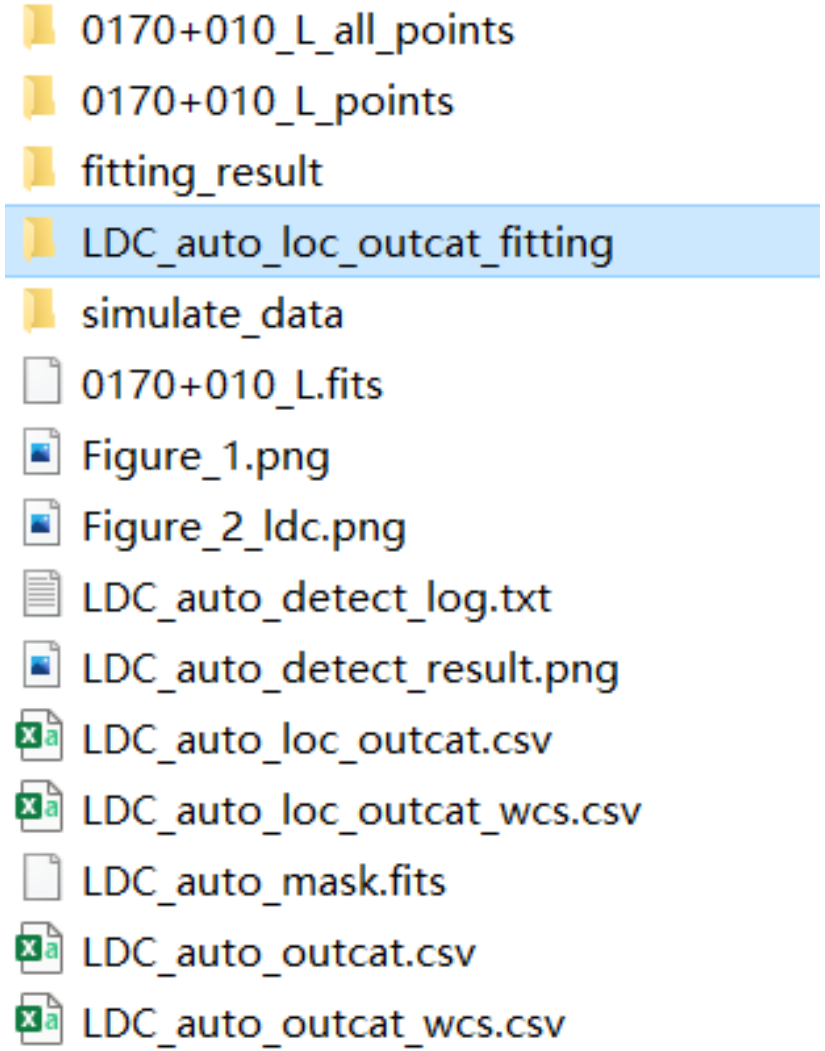
ID	Peak1	Peak2	Peak3	Cen1	Cen2	Cen3	Size1	Size2	Size3	theta	Peak	Sum	Volume
0.00000	51.00000	51.00000	51.00000	51.00000	51.00000	51.00000	14.12900	7.06400	11.77400	36.00000	3.80000	5386.36700	15455.00000
1.00000	41.00000	47.00000	43.00000	41.00000	47.00000	43.00000	14.12900	7.06400	11.77400	36.00000	3.80000	5386.36700	15455.00000

仿真云核的参数

ID	Peak1	Peak2	Peak3	Cen1	Cen2	Cen3	Size1	Size2	Size3	theta	Peak	Sum	success	cost
0	50.99999	51.00001	50.99999	50.99999	51.00001	50.99999	7.06442	14.12840	11.77401	35.99999	3.80001	5386.24014	True	0.00000
1	41.00001	47.00000	43.00001	41.00001	47.00000	43.00001	7.06442	14.12840	11.77401	35.99999	3.80001	5386.24091	True	0.00000

拟合的核表

选取MWISP分子云数据，0170+010 的数据块，基于LDC检测结果进行试验

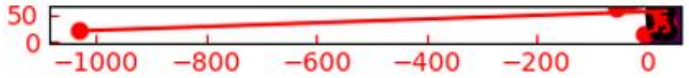


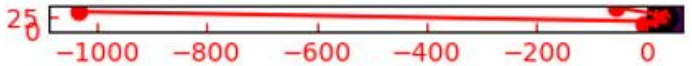
Data information
data file: /home/data/clumps_share/MWISP/R2_200/0170+010/0170+010_L.fits
the rms of data: 0.25320
data shape: [2411 121 121]
==========

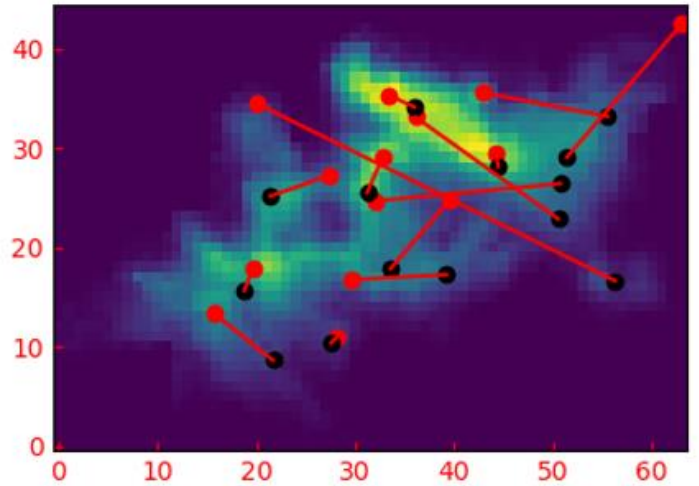
Algorithm parameter information
rho_min[3*rms] = 1.266
delta_min[4] = 4.000
v_min[27] = 27.000
gradmin[0.01] = 0.010
noise[2*rms] = 0.506
dc = None
==========

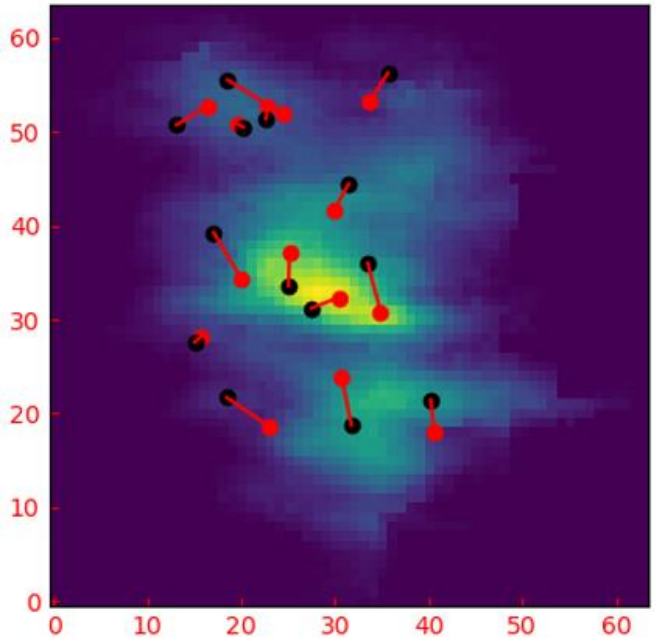
Detect result
69 clumps are rejected because they touched the border.
The number of clumps: 234
The number of local region clumps: 85
delata, rho and Gradient are calculated, using 205.81 seconds.
Outcats are calculated, using 1384.48 seconds.

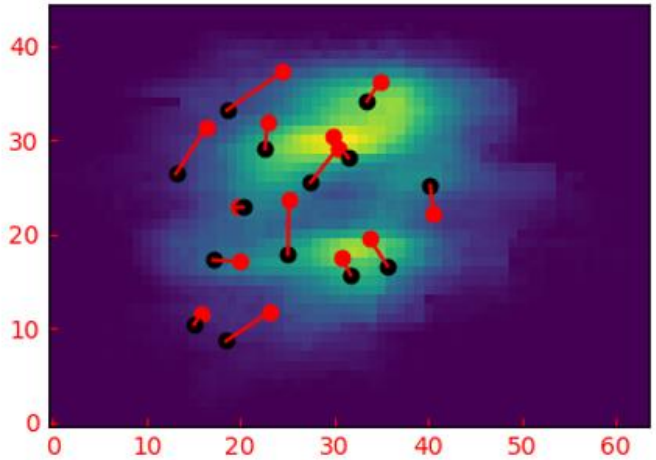
多高斯成分拟合在真实分子云数据上的实验

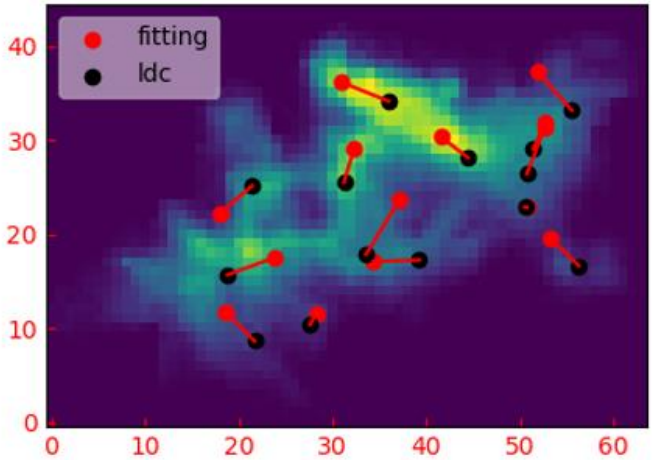




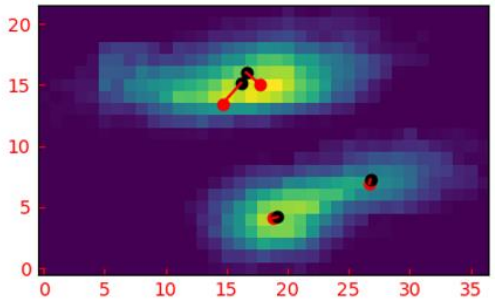


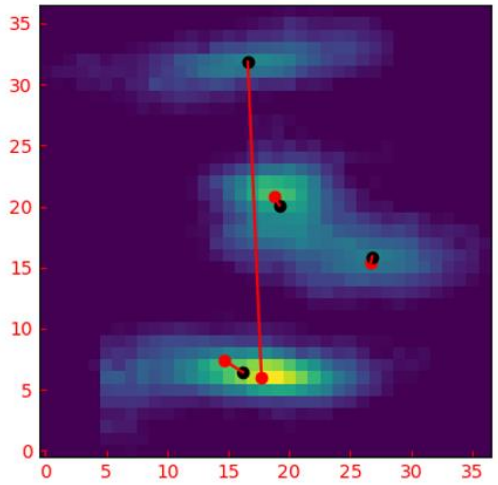


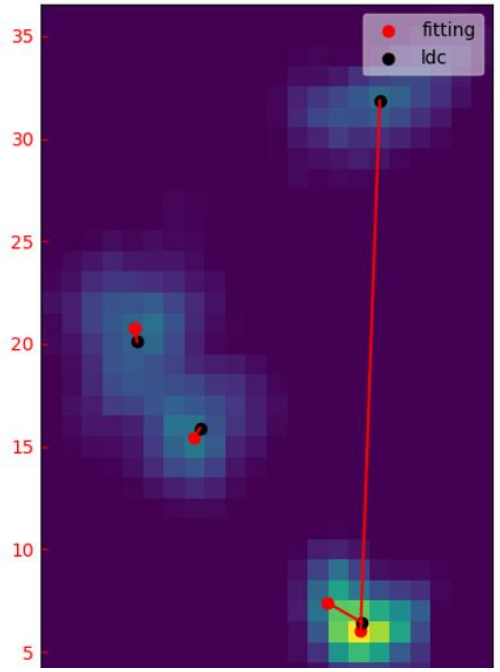




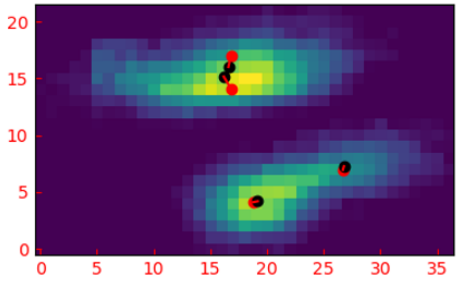
没有对拟合参数做限制的拟合结果

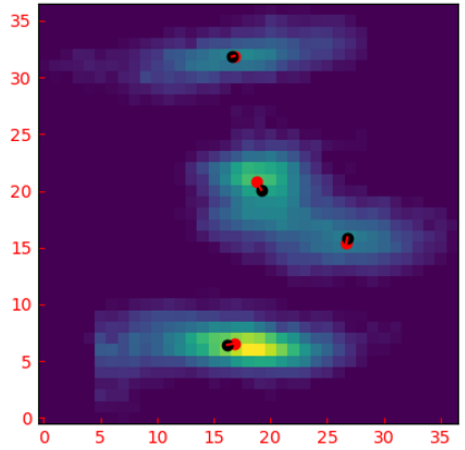


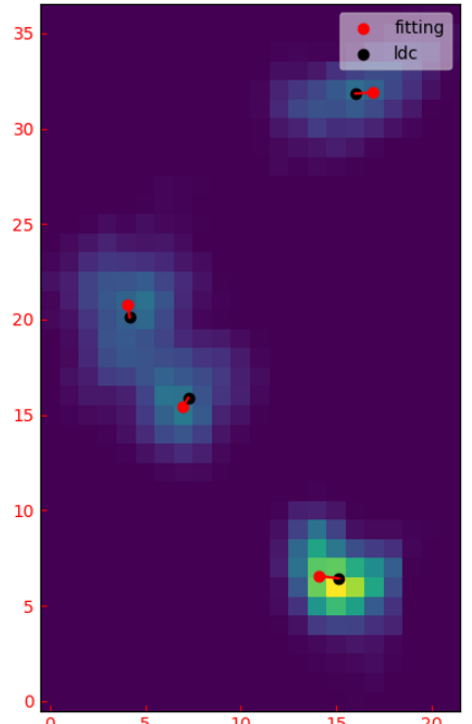


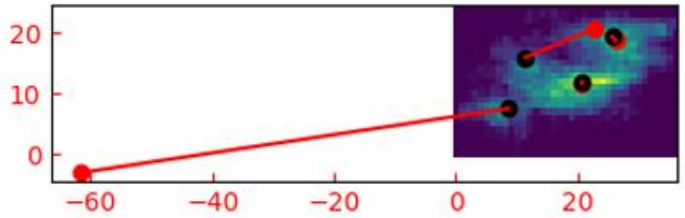


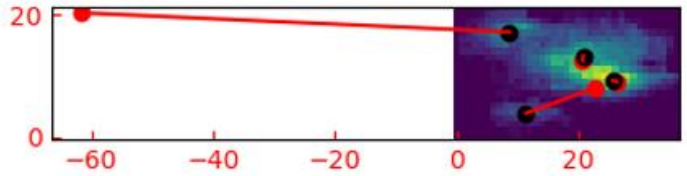
对拟合参数做限制的拟合结果

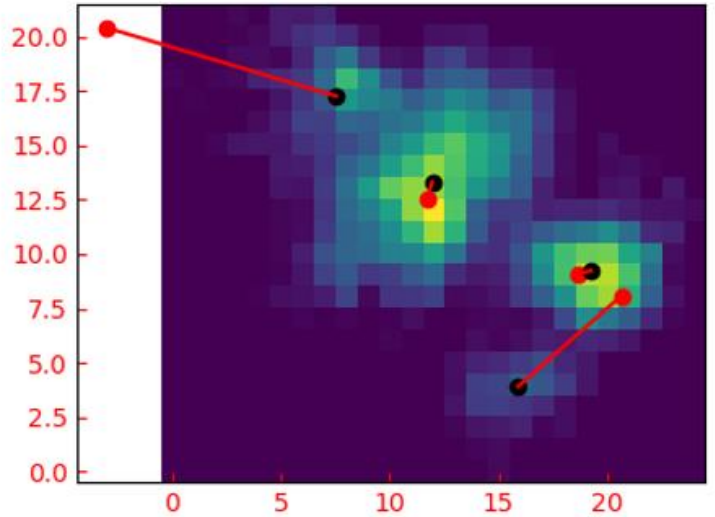


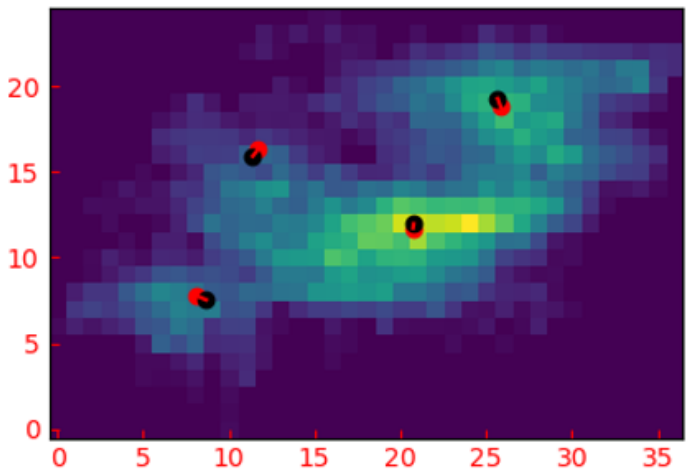


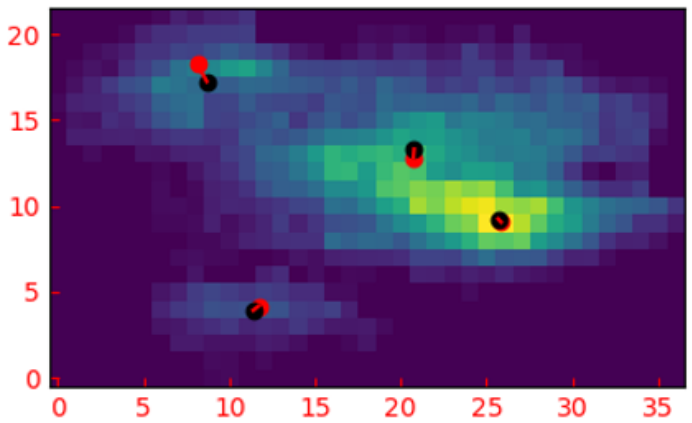


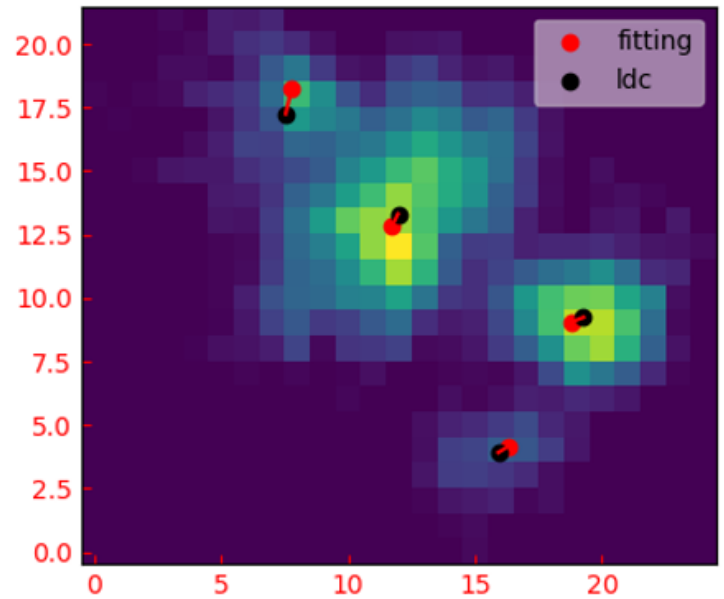

20
10
0
−60
−40
−20
0
20


20
0
−60
−40
−20
0
20

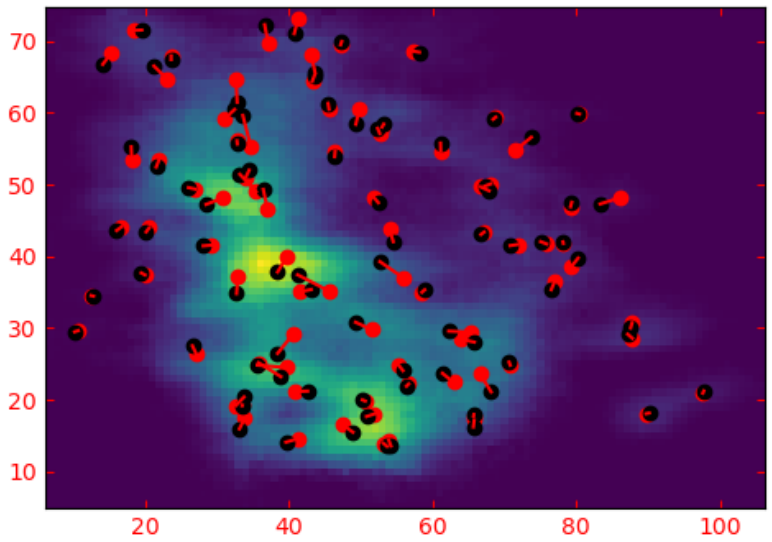

20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0
5
10
15
20

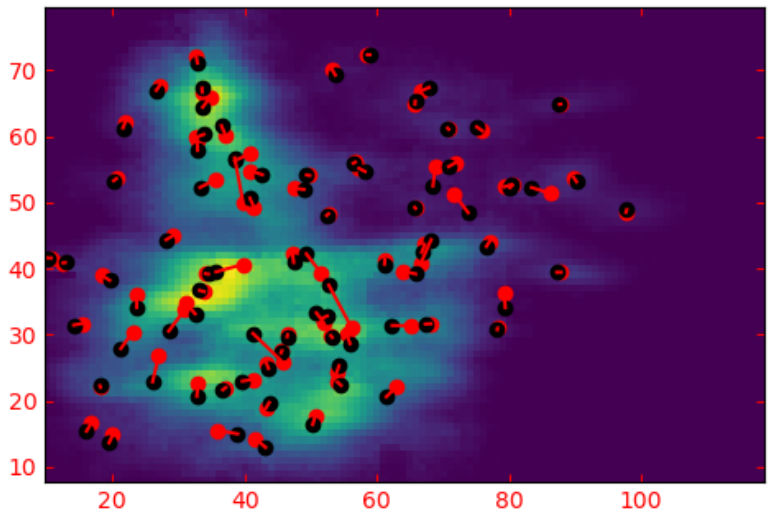

20
15
10
5
0
0
5
10
15
20
25
30
35

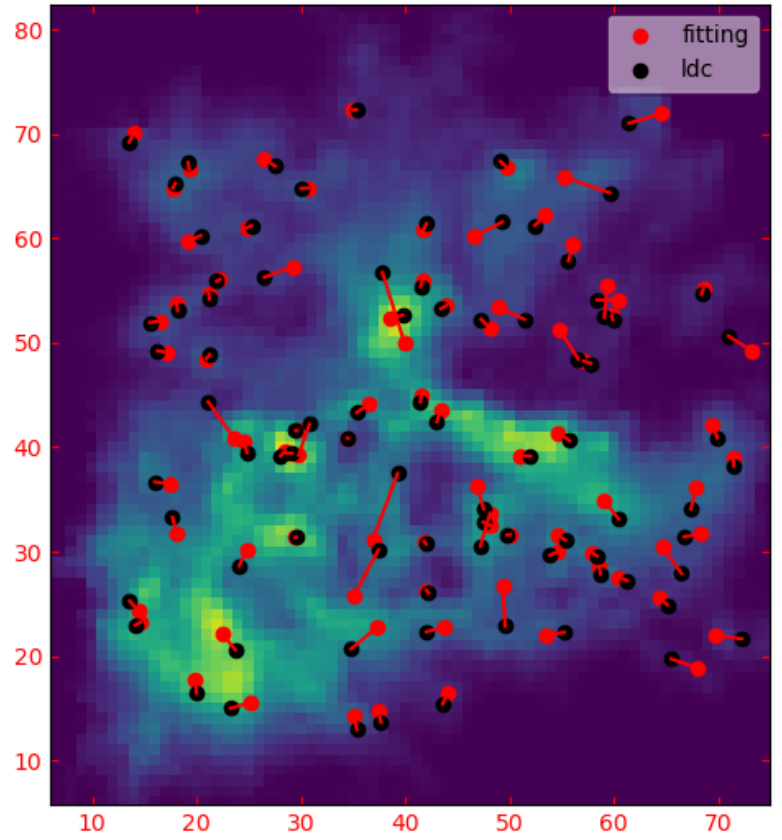

20
15
10
5
0
0
5
10
15
20
25
30
35


fitting
ldc
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0
5
10
15
20

所有拟合结果前后对比，在三个方向的积分图上的展示

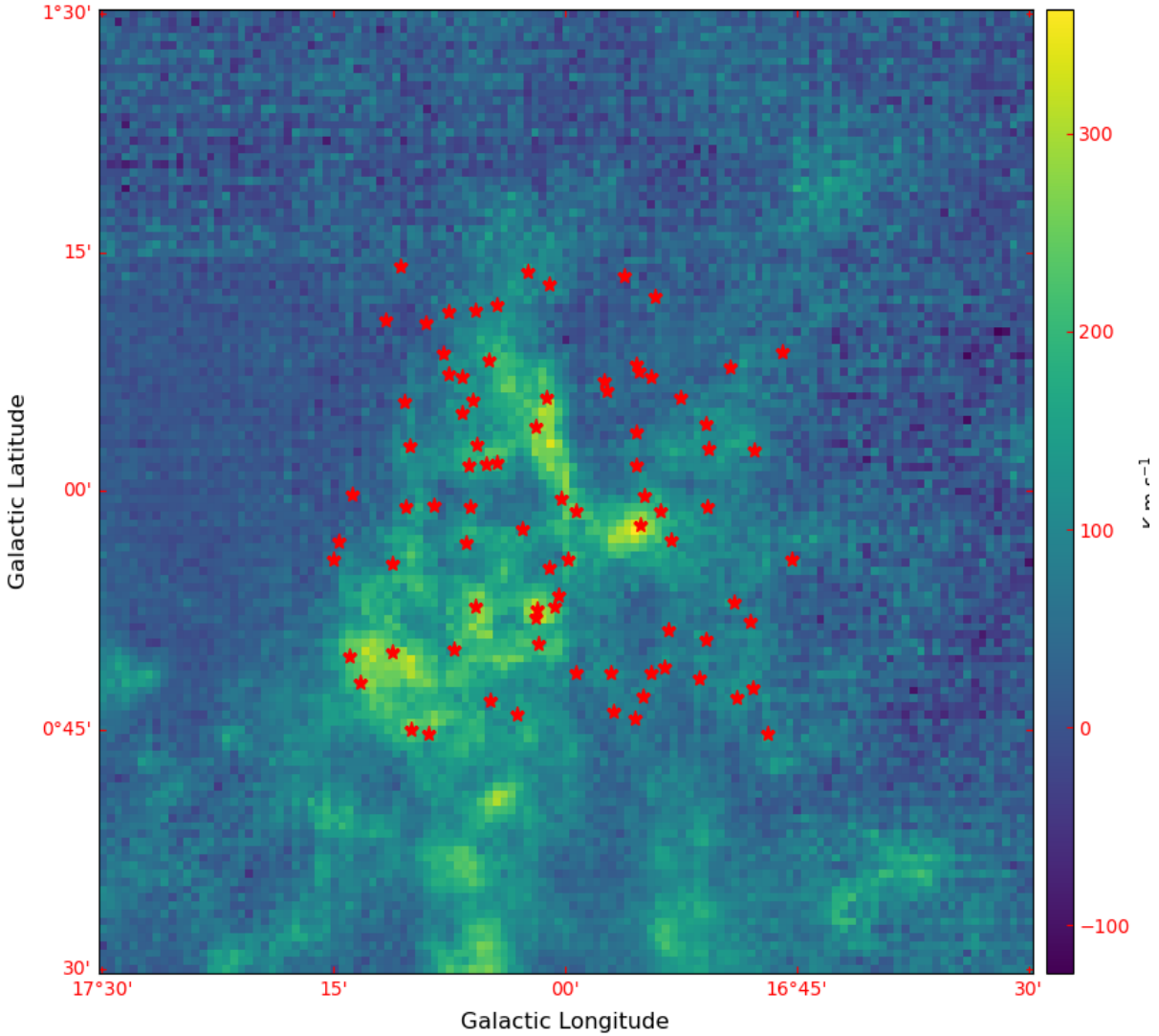






本次试验的cell文件拟合完成，用时约1小时

原始结果



拟合结果