

Diatrack Session File Guide

Johns Hopkins University, Wu Lab

Author: Sun Jay Yoo

Data: June 26, 2017

The Diatrack session file is stored on a .mat (Microsoft Access Table) file that is used by MATLAB as a container for functions and variables. The exported Diatrack .mat session file contains many different variables including color maps, dimensions, indexes, images, and tracks. For analysis purposes, only the *tracks* variable is of interest. In .mat files, it is also important to note that variable indexing begins at 1.

While it may seem as though the *tracks* variable stores information indexed by the different single molecule trajectories, it actually stores information indexed by frame. The index length of the *tracks* variable should be equivalent to the number of frames in the file. For this example, this session file has 10117 frames:

```
>> length(tracks)

ans =

    10117
```

Each index of *tracks* is a structure that contains information about every molecule that appears in the frame. For each molecule in that frame, *tracks* records the X, Y, Z coordinates, intensity, goodness of fit, successor, and predecessor. This is the first index of the *tracks* variable, the first frame, and it has 5 molecules:

```
>> tracks{1}

ans =

    struct with fields:

    RefinedCooY: [11.7849 44.1734 75.9684 73.8330 60.6965]
    RefinedCooX: [38.8411 60.0523 64.5863 74.8155 82.8679]
    RefinedCooZ: [1.0001 1.0001 1.0000 1.0001 1.0001]
    Intensity: [126.4305 124.8418 92.7616 129.8724 230.6995]
    GoodnessFit: [6.4093e-06 7.8323e-06 2.4948e-06 3.8869e-06 3.8296e-06]
    Successor: [0 2 3 5 6]
    Predecessor: [0 0 0 0 0]
```

In order to extract the trajectories of the molecules across different frames, we are only interested in the *RefinedCooY*, *RefinedCooX*, *RefinedCooZ*, *Successor*, and *Predecessor* variables for each molecule in the frame. It is highly important to note that *Predecessor* is usually always listed before *Successor* as the sixth and seventh variable in each frame. Diatrack, for an unknown reason, seems to reverse this order for the first frame only. Molecule information is stored by index in each frame. The first entries refer to one variable, the second entries refer to another, and so on.

RefinedCooY, *RefinedCooX*, and *RefinedCooZ* refer to the Y, X, and Z coordinates of the molecule in frame (ordered in that manner at all times), and is stored numerically (from smallest

to largest) based on the X coordinate value for each frame. *Successor* refers to the index in the next frame in which this molecule appears and *Predecessor* refers to the index in the previous frame. For the first frame or first index of *tracks*, every molecule should have a value of 0 for *Predecessor* as there is no previous frame. In our example above, the first molecule listed has a *Successor* of 0, meaning that it is no longer seen in the next frame, while the next four molecules have successors that appear in indexes 2, 3, 5, and 6 in the next frame. The second index of *tracks*, the second frame, has six molecules:

```
>> tracks{2}

ans =

struct with fields:

  RefinedCooY: [37.3711 43.6777 76.0090 70.4019 74.6256 60.8150]
  RefinedCooX: [29.0089 59.0154 64.7354 73.0693 75.0821 82.8030]
  RefinedCooZ: [1.0001 1.0000 1.0001 1.0000 1.0000 1.0001]
  Intensity: [84.8280 106.2075 144.0688 133.8777 128.6550 232.1278]
  GoodnessFit: [5.7017e-06 3.8591e-06 4.1910e-06 9.8136e-06 2.1817e-06 5.2488e-06]
  Predecessor: [0 2 3 0 4 5]
  Successor: [0 1 2 3 0 4]
```

As expected, the second, third, fifth, and sixth molecule recorded in this frame refer to the corresponding molecule in the previous frame (a trajectory). *Predecessor* data recorded in this second frame also show that these molecules did indeed come from the second, third, forth, and fifth molecules of the previous frame. The first molecule in this second frame has *Predecessor* and *Successor* values of 0, meaning that this is the first frame that this molecule appears in and it no longer exists in the next frame. The fifth molecule in this second frame has a *Successor* value of 0, meaning it no longer appears in the third frame- thus being a trajectory across two frames. This diagram shows an example of how the trajectory of molecules can be tracked in parallel:

