```
In [20]:  import numpy as np
          import matplotlib.pyplot as plt
          from numpy import genfromtxt
```

```
In [21]:  folder = '/Users/mpiekenbrock/WaCS/dbscan/ignore/hdbscan_ex.csv'
          test_data = genfromtxt(folder, delimiter=',')
```

```
In [22]:  np.size(test_data[:, 1])
```

Out[22]: 100

```
In [23]:  import hdbscan
```

```
In [24]:  clusterer = hdbscan.HDBSCAN(min_cluster_size=5, gen_min_span_tree=True,
                                      approx_min_span_tree=False)
          clusterer.fit(test_data)
```

```
Out[24]: HDBSCAN(algorithm='best', allow_single_cluster=False, alpha=1.0,
             approx_min_span_tree=False, core_dist_n_jobs=4, gen_min_span_tree=T
         rue,
             leaf_size=40, memory=Memory(cachedir=None), metric='euclidean',
             min_cluster_size=5, min_samples=None, p=None)
```

```
In [25]:  clusterer.minimum_spanning_tree_.plot(edge_cmap='viridis',
                                                edge_alpha=0.6,
                                                node_size=80,
                                                edge_linewidth=2)
```

Out[25]: <matplotlib.axes._subplots.AxesSubplot at 0x117045f50>

```
In [26]:  # clusterer.minimum_spanning_tree_.to_numpy()
```

```
In [27]:  np.sum(clusterer.minimum_spanning_tree_.to_numpy()[:, 2])
```

Out[27]: 33.506381488456363