WEBVTT

1 00:00:00.210 --> 00:00:02.370 <v ->Today it's my pleasure to introduce,</v>

2 00:00:02.370 --> 00:00:04.560 Professor Ali Shojaie.

3 00:00:04.560 --> 00:00:07.260 Professor Shojaie holds master's degrees

4 00:00:07.260 --> 00:00:09.630 in industrial engineering, statistics,

5 00:00:09.630 --> 00:00:12.570 applied math, and human genetics.

6 00:00:12.570 --> 00:00:14.460 He earned his PhD in statistics

7 00:00:14.460 --> 00:00:16.680 from the University of Michigan.

8 00:00:16.680 --> 00:00:19.230 His research focuses on the high dimensional data,

9 00:00:19.230 --> 00:00:23.160 longitudinal data, computational biology,

10 00:00:23.160 --> 00:00:26.310 network analysis, and neuroimaging.

11 00:00:26.310 --> 00:00:29.070 Professor Shojaie is a 2022 fellow

12 00:00:29.070 --> 00:00:31.590 of the American Statistical Association

13 00:00:31.590 --> 00:00:36.210 and 2022 winner of their Leo Breiman Award.

14 00:00:36.210 --> 00:00:38.280 He's a full professor of biostatistics,

15 00:00:38.280 --> 00:00:40.230 adjunct professor of statistics,

16 00:00:40.230 --> 00:00:43.380 and the associate chair for strategic research affairs

17 00:00:43.380 --> 00:00:44.970 in the department of biostatistics

18 00:00:44.970 --> 00:00:46.980 in the University of Washington.

19 00:00:46.980 --> 00:00:48.580 Let's welcome Professor Shojaie.

20 00:00:51.750 --> 00:00:52.900 <v ->Thanks for having me.</v>

21 00:00:53.760 --> 00:00:57.450 Sometimes I get moved by the volume of my voice.

22 00:00:57.450 --> 00:00:59.730 You guys, can you hear me at the back, okay?

23 00:00:59.730 --> 00:01:01.494 Since I'm not gonna use the microphone yet,

24 00:01:01.494 --> 00:01:04.503 but I'd rather not use the microphone at all.

25 00:01:05.850 --> 00:01:08.250 Well, it's a pleasure to be here

26 00:01:08.250 --> 00:01:11.937 and to talk to you about some work that I've doing doing

27 00:01:11.937 --> 00:01:13.563 for the past couple of years.

28 00:01:14.880 --> 00:01:19.780 I'm using machine learning tools for different types of data

29 00:01:20.785 --> 00:01:25.785 that you can understand better how the brain works.

30 00:01:28.800 --> 00:01:32.330 The question really is how do we process

31 00:01:32.330 --> 00:01:34.290 information on our brains?

32 00:01:34.290 --> 00:01:37.383 What is the processing information?

33 00:01:40.620 --> 00:01:42.810 The brain through neurons,

34 00:01:42.810 --> 00:01:45.900 we know that neurons interact with each other.

35 00:01:45.900 --> 00:01:48.150 Neurons do process information.

36 00:01:51.324 --> 00:01:53.910 This is of course related to my broader interests

37 00:01:53.910 --> 00:01:57.390 on network and understanding how things interact

38 00:01:57.390 --> 00:01:58.620 with each other.

39 00:01:58.620 --> 00:02:02.658 Naturally I was drawn into this part here,

40 00:02:02.658 --> 00:02:05.781 but when I talk to scientist colleagues,

41 00:02:05.781 --> 00:02:07.590 then a lot of times I'm asked,

42 00:02:07.590 --> 00:02:09.724 what is the goal of understanding that network?

43 00:02:09.724 --> 00:02:10.557 How do we use it?

44 00:02:10.557 --> 00:02:11.390 How do we

45 00:02:15.037 --> 00:02:17.400 take advantage of that network that we learned?

46 00:02:17.400 --> 00:02:21.360 Here's an example of some recent work that we've been doing

47 00:02:21.360 --> 00:02:26.280 that indicates that learning something about these networks

48 00:02:26.280 --> 00:02:28.233 is actually important.

49 00:02:30.090 --> 00:02:31.638 I should say that this is joint work

50 00:02:31.638 --> 00:02:36.220 with a bunch of colleagues at the University of Washington

51 00:02:38.100 --> 00:02:41.583 has done that is biomedical engineering,

52 00:02:42.640 --> 00:02:46.620 and the main group that has been running these experiments.

53 00:02:46.620 --> 00:02:49.290 And then I'm collaborating with E Shea-Brown

54 00:02:49.290 --> 00:02:51.098 who's in computational scientist,

55 00:02:51.098 --> 00:02:55.173 and Z Harchaoui, computer scientist slash statistician,

56 00:02:56.010 --> 00:02:58.713 and she's been working on this project.

57 00:02:58.713 --> 00:03:01.560 This project, the lab is interested.

58 00:03:01.560 --> 00:03:05.070 And what they do is neurostimulation.

59 00:03:05.070 --> 00:03:08.220 What they wanna do is to see if they could stimulate

60 00:03:08.220 --> 00:03:12.120 in different regions of the brain to make in this case

61 00:03:12.120 --> 00:03:13.590 monkey do certain things

62 00:03:13.590 --> 00:03:17.373 or to restore function that the monkey might have lost.

63 00:03:18.210 --> 00:03:22.110 And it's a really interesting platform

64 00:03:22.110 --> 00:03:23.260 that they've developed.

65 00:03:24.360 --> 00:03:27.960 It's basically small implants that they put

66 00:03:27.960 --> 00:03:31.273 in a region of the brain on these monkeys.

67 00:03:31.273 --> 00:03:35.490 And the implant has two areas when the lasers

68 00:03:35.490 --> 00:03:40.490 beam shine in about 96 in this case,

69 00:03:40.710 --> 00:03:42.520 electrodes that collect data

70 00:03:43.476 --> 00:03:45.176 in that small region of the brain.

71 00:03:46.590 --> 00:03:50.790 This is made possible by optogenetics

72 00:03:50.790 --> 00:03:54.960 meaning that it made the neurons sensitive to these lasers.

73 00:03:54.960 --> 00:03:56.440 When neurons

74 00:03:59.610 --> 00:04:02.520 receive the laser, then they basically get excited,

75 00:04:02.520 --> 00:04:03.933 get activate.

76 00:04:04.950 --> 00:04:07.560 The goal in this research eventually

77 00:04:07.560 --> 00:04:09.933 is to see how the activation of neurons,

78 00:04:10.890 --> 00:04:14.490 which plasticity would change

79 00:04:14.490 --> 00:04:16.090 the connectivity of the neurons,

80 00:04:18.360 --> 00:04:22.560 would result in later on in changing function.

81 00:04:22.560 --> 00:04:24.270 That's the eventual goal of this.

82 00:04:24.270 --> 00:04:28.290 This research work at the very beginning of that.

83 00:04:28.290 --> 00:04:31.650 We are not there yet in terms of understanding function,

84 00:04:31.650 --> 00:04:34.530 understanding the link, the connectivity and contact.

85 00:04:34.530 --> 00:04:37.440 The collaboration with this lab started

86 00:04:37.440 --> 00:04:41.070 when they wanted to predict how the connectivity changes

87 00:04:41.070 --> 00:04:43.263 as a result of this activation.

88 00:04:44.190 --> 00:04:48.737 We wanted to understand whether by changing various factors

89 00:04:48.737 --> 00:04:52.020 in the experiments, the distance between two lasers,

90 00:04:52.020 --> 00:04:53.970 the duration of laser.

91 00:04:53.970 --> 00:04:57.723 How could they accurately predict the changing connectivity?

92 00:05:00.912 --> 00:05:02.010 The way that the experiment is set up

93 00:05:02.010 --> 00:05:06.130 is that basically had these times where they have

94 00:05:07.290 --> 00:05:09.990 activation and then the latency period

95 00:05:09.990 --> 00:05:11.670 and then followed by observation.

96 00:05:11.670 --> 00:05:16.020 They basically observe the activity of these brain regions.

97 00:05:19.560 --> 00:05:20.853 That sort of 96.

98 00:05:22.350 --> 00:05:25.380 Electrodes in this main region over time.

99 00:05:25.380 --> 00:05:27.230 That's the data that they're correct.

100 00:05:30.930 --> 00:05:34.920 Here's a look at this functional connectivity a

101 00:05:34.920 --> 00:05:38.373 and that's what they were trying to predict.

102 00:05:39.510 --> 00:05:42.880 Basically the heat map shows

103 00:05:46.061 --> 00:05:49.500 the links between the various brain lesions,

104 00:05:49.500 --> 00:05:52.633 but 96 of them, you don't wanna.

105 00:05:56.481 --> 00:06:01.320 And if that connectivity is defined based on coherence,

106 00:06:01.320 --> 00:06:04.710 which is basically correlation measure frequency domain,

107 00:06:04.710 --> 00:06:07.890 and we have coherence in four different frequency bands.

108 00:06:07.890 --> 00:06:10.740 These are the standard bands that signal instructive

109 00:06:10.740 --> 00:06:13.800 and they think that they measure activity

110 00:06:13.800 --> 00:06:16.050 and different spatial resolution.

111 00:06:16.050 --> 00:06:18.478 We have theta band, the beta band, the gamma band,

112 00:06:18.478 --> 00:06:20.040 and the high gamma band.

113 00:06:20.040 --> 00:06:22.320 And we wanna see how the connectivity

114 00:06:22.320 --> 00:06:24.510 in these different bands changes

115 00:06:24.510 --> 00:06:26.310 as the effect of these type neurons.

116 00:06:31.184 --> 00:06:32.017 And what...

117 00:06:36.780 --> 00:06:38.430 This is not working.

118 00:06:38.430 --> 00:06:39.750 The clicker stopped working.

119 00:06:39.750 --> 00:06:40.743 We'll figure that.

120 00:06:50.550 --> 00:06:53.200 Let's go on full screen again to see where this goes.

121 00:06:59.790 --> 00:07:01.290 What basically we have

122 00:07:01.290 --> 00:07:03.480 is that we have the baseline connectome

123 00:07:03.480 --> 00:07:06.660 and we have these experimental protocols,

124 00:07:06.660 --> 00:07:10.020 and we're trying to predict how the connectivity changes.

125 00:07:10.020 --> 00:07:11.700 What the lab was doing before was that

126 00:07:11.700 --> 00:07:14.490 they were looking at trying to predict connectivity

127 00:07:14.490 --> 00:07:18.240 based on experimental protocols.

128 00:07:18.240 --> 00:07:19.320 And what they were getting

129 00:07:19.320 --> 00:07:22.410 was actually really bad prediction.

130 00:07:22.410 --> 00:07:25.800 These are test R squares.

131 00:07:25.800 --> 00:07:29.700 And what they were getting was about 5% test R square

132 00:07:29.700 --> 00:07:31.620 when they were using these protocol features

133 00:07:31.620 --> 00:07:34.470 to predict how to connect with these gene.

134 00:07:34.470 --> 00:07:35.760 And the first thing that we understood

135 00:07:35.760 --> 00:07:38.250 and so you see it that sort of really bad

136 00:07:38.250 --> 00:07:39.330 is that that's the prediction.

137 00:07:39.330 --> 00:07:40.710 If that's the prediction that you're getting,

138 00:07:40.710 --> 00:07:42.183 then really bad prediction.

139 00:07:43.320 --> 00:07:45.537 The first thing that we noticed in this research

140 00:07:45.537 --> 00:07:49.560 was that it's actually important to incorporate

141 00:07:49.560 --> 00:07:52.740 the features of the current state of connectivity

142 00:07:52.740 --> 00:07:54.940 in order to predict how to make them useful.

143 00:07:56.340 --> 00:07:59.430 What we did was that in addition to those protocol features,

144 00:07:59.430 --> 00:08:01.380 we added some network features,

145 00:08:01.380 --> 00:08:03.390 the current state of the network in order to predict

146 00:08:03.390 --> 00:08:04.440 how it's gonna change.

147 00:08:04.440 --> 00:08:06.240 And this is, to me, this is really interesting

148 00:08:06.240 --> 00:08:09.660 because it basically says that our prediction

149 00:08:09.660 --> 00:08:12.570 has to be subject specific

150 00:08:12.570 --> 00:08:13.982 depending on the current state of each month

151 00:08:13.982 --> 00:08:17.790 these connectivity, how their connectivity

152 00:08:17.790 --> 00:08:19.923 is going to change will be different.

153 00:08:20.820 --> 00:08:24.060 And what we saw was that when we incorporated

154 00:08:24.060 --> 00:08:27.660 these network features, we were able to improve quite a bit

155 00:08:27.660 --> 00:08:28.680 in terms of prediction.

156 00:08:28.680 --> 00:08:33.180 We're still not doing hugely good,

157 00:08:33.180 --> 00:08:36.300 we're only getting like test R squared of what, 25%.

158 00:08:36.300 --> 00:08:38.190 But what you see that sort of the connectivity

159 00:08:38.190 --> 00:08:40.974 is now, the prediction is now much more.

160 00:08:40.974 --> 00:08:42.925 How the connectivity.

161 00:08:42.925 --> 00:08:46.440 And also in terms of the pictures, you see that going from,

162 00:08:46.440 --> 00:08:48.360 so say this is the true,

163 00:08:48.360 --> 00:08:51.600 the first part in d is the true change in connectivity,

164 00:08:51.600 --> 00:08:55.620 e is what you would get from just the protocol features,

165 00:08:55.620 --> 00:08:57.250 and you see that prediction is really bad,

166 00:08:57.250 --> 00:09:00.510 and f is what you get when you combine protocol features

167 00:09:00.510 --> 00:09:02.133 and the network features.

168 00:09:03.360 --> 00:09:05.950 That prediction is closer to the true

169 00:09:08.550 --> 00:09:12.420 change in connectivity than just using the protocol feature.

170 00:09:12.420 --> 00:09:15.180 This was the first thing that we learned from this research.

171 00:09:15.180 --> 00:09:17.760 The second part of what we learned is that

172 00:09:17.760 --> 00:09:20.670 it also matters which approach you used the prediction.

173 00:09:20.670 --> 00:09:24.120 What they had done was that they were using some simple

174 00:09:24.120 --> 00:09:25.560 like linear model for prediction.

175 00:09:25.560 --> 00:09:28.310 And then we realized that we need to use something more

176 00:09:30.000 --> 00:09:32.340 expressive and then we sort of ended up using

177 00:09:32.340 --> 00:09:33.930 these non-linear additive models

178 00:09:33.930 --> 00:09:35.580 that we had previously developed,

179 00:09:35.580 --> 00:09:40.020 partly because while they have a lot of expressive power,

180 00:09:40.020 --> 00:09:42.540 they're still easy to interpret.

181 00:09:42.540 --> 00:09:46.110 Interpretation for these additive models is still easy

182 00:09:46.110 --> 00:09:48.580 and particularly we see what the shapes

183 00:09:50.790 --> 00:09:52.170 basically these functions are.

184 00:09:52.170 --> 00:09:54.540 For example, with the distance we see how the function

185 00:09:54.540 --> 00:09:57.927 changes and that helps with the design of these experience.

186 00:09:57.927 --> 00:09:59.700 I'm not gonna spend too much time

187 00:09:59.700 --> 00:10:01.170 talking about the details of this

188 00:10:01.170 --> 00:10:03.120 given that we only have 50 minutes

189 00:10:03.120 --> 00:10:04.950 and I wanna get to the main topic,

190 00:10:04.950 --> 00:10:08.220 but basically these additive models

191 00:10:08.220 --> 00:10:10.800 are built by combining these features.

192 00:10:10.800 --> 00:10:14.250 Think of tailor expansion in a very simple sense

193 00:10:14.250 --> 00:10:17.010 that you have a linear term, you have a quadratic term,

194 00:10:17.010 --> 00:10:18.180 you have a cubic term.

195 00:10:18.180 --> 00:10:21.270 And the way that sort we form these additive models

196 00:10:21.270 --> 00:10:25.650 is that we automatically select the degree of complexity

197 00:10:25.650 --> 00:10:27.960 of each additive feature,

198 00:10:27.960 --> 00:10:32.370 whether it's says linear, or quadratic, or cubic, etcetera.

199 00:10:32.370 --> 00:10:36.210 We also allow some features to be present in the models,

200 00:10:36.210 --> 00:10:37.470 features not to be present.

201 00:10:37.470 --> 00:10:40.710 What we end up with are these patterns

202 00:10:40.710 --> 00:10:43.050 where some features are real complex and other features,

203 00:10:43.050 --> 00:10:45.200 and that's automatically decided from data.

204 00:10:46.950 --> 00:10:50.940 This model is good in this prediction

205 00:10:50.940 --> 00:10:53.310 and it allows us to come up with these sets of predictions.

206 00:10:53.310 --> 00:10:57.507 We see now that for example, for coherence difference,

207 00:10:57.507 --> 00:10:59.250 which is the network feature,

208 00:10:59.250 --> 00:11:01.200 that's the coherence difference.

209 00:11:01.200 --> 00:11:02.730 Network distance, that's the distance

210 00:11:02.730 --> 00:11:03.660 between the two portals.

211 00:11:03.660 --> 00:11:05.160 The two laser points.

212 00:11:05.160 --> 00:11:07.410 We get these two patterns estimated

213 00:11:07.410 --> 00:11:10.350 and then when we combine them, we get this surface basically

214 00:11:10.350 --> 00:11:15.240 that determines how the connectivity,

215 00:11:15.240 --> 00:11:16.800 changing connectivity could be predicted

216 00:11:16.800 --> 00:11:17.670 based on these two features.

217 00:11:17.670 --> 00:11:21.603 And all of this is done automatically based on data.

218 00:11:22.860 --> 00:11:24.930 This approach, again, sort of the key feature of it

219 00:11:24.930 --> 00:11:27.930 is that it combines the network features

220 00:11:27.930 --> 00:11:29.900 of the current state of connectivity with protocol features

221 00:11:29.900 --> 00:11:32.880 in order to do a better job of prediction.

222 00:11:32.880 --> 00:11:36.240 This is a research that we just started

223 00:11:36.240 --> 00:11:39.120 and we will continue this research

224 00:11:39.120 --> 00:11:40.770 for the next at least five years.

225 00:11:42.352 --> 00:11:43.946 But the goal of it is eventually to see

226 00:11:43.946 --> 00:11:46.340 if we could predict the function

227 00:11:46.340 --> 00:11:48.540 and ultimately if we could build a controller

228 00:11:48.540 --> 00:11:51.570 that we could determine how to change function

229 00:11:51.570 --> 00:11:54.783 based on various features of the experiment.

230 00:11:57.230 --> 00:11:59.250 I mentioned all of this to say that knowing

231 00:11:59.250 --> 00:12:01.230 and learning the network matters.

232 00:12:01.230 --> 00:12:03.780 We need to learn the current state of connectivity,

233 00:12:03.780 --> 00:12:06.930 for example, in this work in order to be able to design

234 00:12:06.930 --> 00:12:09.247 experiments that would hopefully help

235 00:12:12.030 --> 00:12:14.850 and restore function.

236 00:12:14.850 --> 00:12:17.340 Now in this particular work,

237 00:12:17.340 --> 00:12:19.950 what we did was that we used a very simple

238 00:12:19.950 --> 00:12:20.940 notion of connectivity.

239 00:12:20.940 --> 00:12:23.910 We used coherence, which is basically correlation,

240 00:12:23.910 --> 00:12:26.980 but we know that that’s not always the best

241 00:12:28.110 --> 00:12:32.460 way to define connectivity between ranges.

242 00:12:32.460 --> 00:12:35.970 And so what I wanna talk about for the remaining

243 00:12:35.970 --> 00:12:40.080 40 minutes or so is how do we learn connectivity

244 00:12:40.080 --> 00:12:41.790 between neurons?

245 00:12:41.790 --> 00:12:44.820 And this is using a different type of data

246 00:12:44.820 --> 00:12:46.170 that I had thought about before,

247 00:12:46.170 --> 00:12:48.670 and I’m hoping that so I could show you this clip,

248 00:12:51.390 --> 00:12:54.777 which is that shows the actual raw data.

249 00:12:54.777 --> 00:12:56.703 The data is actually a video.

250 00:12:57.660 --> 00:12:59.673 And this is activity of individual neurons

251 00:12:59.673 --> 00:13:02.850 in a small region of the brain.

252 00:13:02.850 --> 00:13:04.207 These dots that you see popping up,

253 00:13:04.207 --> 00:13:07.923 these are individual neurons firing over time.

254 00:13:10.395 --> 00:13:11.970 And you see that sort of neuron fires

255 00:13:11.970 --> 00:13:15.420 and other neuron fires, et cetera, et cetera.

256 00:13:15.420 --> 00:13:17.550 That’s the raw data that we’re getting.

257 00:13:17.550 --> 00:13:21.060 And the goal is to understand

258 00:13:21.060 --> 00:13:23.520 based on this pattern of activation of neurons,

259 00:13:23.520 --> 00:13:26.640 how neurons talk to each other basically.

260 00:13:26.640 --> 00:13:28.173 Now I'm gonna go back here.

261 00:13:34.317 --> 00:13:37.590 And so the data of that video that I showed you,

262 00:13:37.590 --> 00:13:40.920 basically, here's some snapshot of that data.

263 00:13:40.920 --> 00:13:43.047 Here's one frame.

264 00:13:43.047 --> 00:13:46.200 And there's a lot of steps in getting this data

265 00:13:46.200 --> 00:13:48.243 to place it a bit more quick.

266 00:13:49.614 --> 00:13:50.970 Were not gonna talk about this,

267 00:13:51.807 --> 00:13:54.990 but sort of we need to first identify where the neurons are.

268 00:13:54.990 --> 00:13:57.780 No one tells us where the neurons are in that video.

269 00:13:57.780 --> 00:13:59.880 We need to first identify where the neurons are.

270 00:13:59.880 --> 00:14:03.150 We need to identify when they swipe, when they fire.

271 00:14:03.150 --> 00:14:04.950 No one tells us that either.

272 00:14:04.950 --> 00:14:08.700 There's a lot of pre processing step that happens.

273 00:14:08.700 --> 00:14:10.680 The first task is called segmentation,

274 00:14:10.680 --> 00:14:12.510 identifying where the neurons are,

275 00:14:12.510 --> 00:14:15.300 then spike detection, when the nuance fire over time,

276 00:14:15.300 --> 00:14:17.130 when which individual neuron fires over time.

277 00:14:17.130 --> 00:14:19.200 And that none of these is a trivial task.

278 00:14:19.200 --> 00:14:22.318 And then a lot of smart people are working on these,

279 00:14:22.318 --> 00:14:24.600 including some of my colleagues.

280 00:14:24.600 --> 00:14:26.460 After a lot of pre-processing,

281 00:14:26.460 --> 00:14:27.960 so you end up with each individual neuron,

282 00:14:27.960 --> 00:14:31.260 you end up with a data point, like data set like this

283 00:14:31.260 --> 00:14:35.400 that it basically has these takes

284 00:14:35.400 --> 00:14:36.900 whenever the neuron has fired.

285 00:14:39.180 --> 00:14:42.120 A given neuron you have over time that the neuron fire

286 00:14:42.120 --> 00:14:43.953 like this.

287 00:14:45.011 --> 00:14:47.280 These are the time points the neuron apply.

288 00:14:47.280 --> 00:14:48.840 Now, you can do something fancier,

289 00:14:48.840 --> 00:14:51.210 you can look at the magnitude,

290 00:14:51.210 --> 00:14:53.310 the signal that you're detecting at neuron.

291 00:14:53.310 --> 00:14:55.470 You could deal with that, but for now we're ignoring that.

292 00:14:55.470 --> 00:14:57.900 We're just looking at when they fire.

293 00:14:57.900 --> 00:15:00.053 This is called the spike train for each neuron.

294 00:15:01.200 --> 00:15:03.423 That's the data that we're using.

295 00:15:04.507 --> 00:15:07.080 These are neurons firing times.

296 00:15:07.080 --> 00:15:09.120 And if we combine them, this is the cartoon

297 00:15:09.120 --> 00:15:09.953 we get something like this.

298 00:15:09.953 --> 00:15:12.720 We get a sequence of activation pattern.

299 00:15:12.720 --> 00:15:16.230 This is color coded based on that sort of five neuron

300 00:15:16.230 --> 00:15:17.730 sort of cartoon network.

301 00:15:17.730 --> 00:15:19.440 And you see that different neurons activate

302 00:15:19.440 --> 00:15:20.403 at different times.

303 00:15:22.924 --> 00:15:24.870 And what I'll talk about is a notion of connectivity

304 00:15:24.870 --> 00:15:29.130 that tries to predict the activation pattern of one neuron

305 00:15:29.130 --> 00:15:31.170 from a network, basically.

306 00:15:31.170 --> 00:15:33.510 That sort of maybe neuron one tells us something

307 00:15:33.510 --> 00:15:36.120 about sort of activation patterns in neuro two,

308 00:15:36.120 --> 00:15:39.300 that if we knew when neuro one activated or fired,

309 00:15:39.300 --> 00:15:41.370 we could predict when neuro on two fires,

310 00:15:41.370 --> 00:15:43.230 and maybe neuron two will tell us something

311 00:15:43.230 --> 00:15:46.107 about activations of neurons three and four, et cetera.

312 00:15:46.107 --> 00:15:48.600 And that's the notion of connectivity at that time

313 00:15:48.600 --> 00:15:51.390 after, since we're trying to estimate those edges

314 00:15:51.390 --> 00:15:52.830 in this time.

315 00:15:52.830 --> 00:15:54.810 Now, please.

316 00:15:54.810 --> 00:15:56.610 <v ->Could you say just a few words informally</v>

317 00:15:56.610 --> 00:15:58.350 about the direction of connectivity?

318 00:15:58.350 --> 00:15:59.183 <v ->Yeah.</v>

319 00:15:59.183 --> 00:16:00.450 <v ->Maybe drawing arrow forward in time.</v>

320 00:16:00.450 --> 00:16:01.320 <v ->Yes.</v>

321 00:16:01.320 --> 00:16:03.753 I'll get to this, maybe in the next two slides.

322 00:16:05.940 --> 00:16:07.940 The framework that we're gonna work with

323 00:16:08.910 --> 00:16:10.680 is called the Hawkes process.

324 00:16:10.680 --> 00:16:13.980 Just go back to seminal more by Alan Hawkes.

325 00:16:13.980 --> 00:16:18.980 In '70s where he looked at spectral properties

326 00:16:19.140 --> 00:16:20.340 of point processes.

327 00:16:20.340 --> 00:16:22.770 What are point processing that basically is like activation

328 00:16:22.770 --> 00:16:23.603 over time.

329 00:16:23.603 --> 00:16:25.539 Zeros and ones over time.

330 00:16:25.539 --> 00:16:26.943 It could Poisson processes.

331 00:16:28.650 --> 00:16:31.410 What the Hawkes process does in particular

332 00:16:31.410 --> 00:16:36.410 is that it uses the past history of one neuron

333 00:16:37.120 --> 00:16:38.970 to predict the future.

334 00:16:38.970 --> 00:16:41.700 And this goes back to Forest's question

335 00:16:41.700 --> 00:16:44.490 that sort of what is that edge in this case?

336 00:16:44.490 --> 00:16:47.910 This is the notion that is related closely in a special case

337 00:16:47.910 --> 00:16:52.140 of what is known to econometricians as Granger causality

338 00:16:52.140 --> 00:16:55.470 that sort of using past to predict future.

339 00:16:55.470 --> 00:16:57.120 And that's the notion of connectivity

340 00:16:57.120 --> 00:17:02.120 that we're here at, we're after in this particular case.

341 00:17:02.688 --> 00:17:05.310 And what makes this Hawkes process

342 00:17:05.310 --> 00:17:06.930 the convenient for this is that

343 00:17:06.930 --> 00:17:08.490 sort of it's already set up to do this.

344 00:17:08.490 --> 00:17:09.690 I'm gonna present the Hawkes process.

345 00:17:09.690 --> 00:17:13.230 Its simplest form, this is the linear Hawkes process.

346 00:17:13.230 --> 00:17:16.590 And what it is, is that sort o, it's a counting process.

347 00:17:16.590 --> 00:17:19.500 It's just counting the events.

348 00:17:19.500 --> 00:17:24.500 And so that's the event process N.

349 00:17:25.350 --> 00:17:30.350 And that event process has an intensity lambda j

350 00:17:30.600 --> 00:17:33.360 for each neuron is standard i,

351 00:17:33.360 --> 00:17:36.917 which is combination of two terms,

352 00:17:36.917 --> 00:17:40.380 a new I, that's the baseline intensity of that neuron.

353 00:17:40.380 --> 00:17:43.050 That means that if you had nothing else,

354 00:17:43.050 --> 00:17:47.280 this neuron would fire at this rate, but basically random

355 00:17:47.280 --> 00:17:49.180 that would fire at random rate

356 00:17:50.850 --> 00:17:52.740 plus the effect that that neuron

357 00:17:52.740 --> 00:17:54.570 gets from the other neurons.

358 00:17:54.570 --> 00:17:57.213 Every time that there's an activation in neuron,

359 00:17:58.260 --> 00:18:02.610 any neuron j from one to p including neuron i itself,

360 00:18:02.610 --> 00:18:05.127 depending on how long it's been since that activation.

361 00:18:05.127 --> 00:18:07.500 The time it's been, the current time t

362 00:18:07.500 --> 00:18:09.420 and the time of activation of the previous neuron

363 00:18:09.420 --> 00:18:11.070 acquiring or the previous neuron,

364 00:18:11.070 --> 00:18:14.670 some weight function determines how much influence

365 00:18:14.670 --> 00:18:16.830 that neuron pi gets.

366 00:18:16.830 --> 00:18:20.190 This has a flavor of causality,

367 00:18:20.190 --> 00:18:24.330 which is why econometricians call it danger causality.

368 00:18:24.330 --> 00:18:28.740 This is worked by the ranger,

369 00:18:28.740 --> 00:18:30.000 but it's really not causality.

370 00:18:30.000 --> 00:18:31.590 We know that there's beyond,

371 00:18:31.590 --> 00:18:32.940 and so there's a lot of work on this

372 00:18:32.940 --> 00:18:34.173 that's sort, it's only causality

373 00:18:34.173 --> 00:18:36.990 on the day-to-day restrictive assumptions,

374 00:18:36.990 --> 00:18:38.190 talk about in general,

375 00:18:38.190 --> 00:18:40.950 but nonetheless it predicts in the future.

376 00:18:40.950 --> 00:18:42.780 It's a prediction in the future.

377 00:18:42.780 --> 00:18:46.740 And again, sort of in this case this d and i

378 00:18:46.740 --> 00:18:51.740 is our point process, lambda i is our intensity process.

379 00:18:51.930 --> 00:18:53.928 It started itself.

380 00:18:53.928 --> 00:18:56.160 Ui is the background intensity

381 00:18:56.160 --> 00:19:01.160 and tjks are the times when the other neurons

382 00:19:01.350 --> 00:19:02.640 acquired in the past.

383 00:19:02.640 --> 00:19:06.360 And this omega ij is the transfer function.

384 00:19:06.360 --> 00:19:09.180 It determines how much information is passed

385 00:19:09.180 --> 00:19:10.980 from firing your one neuron

386 00:19:10.980 --> 00:19:14.190 to firing of other neurons in the future.

387 00:19:14.190 --> 00:19:16.050 And usually you think that sort of the further

388 00:19:16.050 --> 00:19:19.050 you go in the past, the less information is carrying over.

389 00:19:19.050 --> 00:19:21.150 Usually the types of functions that you consider,

390 00:19:21.150 --> 00:19:23.190 these transfer functions are decay

391 00:19:23.190 --> 00:19:25.020 and how to decay form

392 00:19:25.020 --> 00:19:27.000 that sort of, if you go too far in the past,

393 00:19:27.000 --> 00:19:30.330 there's no information, there's no useful information.

394 00:19:30.330 --> 00:19:33.330 Any question on the basic of this linear Hawkes process

395 00:19:33.330 --> 00:19:38.250 because I'm not gonna present the more complicated version,

396 00:19:38.250 --> 00:19:40.770 but I think this will suffice for our conversation.

397 00:19:40.770 --> 00:19:43.260 I wanna make sure that we're all good

398 00:19:43.260 --> 00:19:44.673 with this simple version.

399 00:19:47.850 --> 00:19:49.893 Okay, so no question on this.

400 00:19:50.910 --> 00:19:54.540 But if we agree with this and then this actually process

401 00:19:54.540 --> 00:19:55.980 gives us a very convenient way

402 00:19:55.980 --> 00:19:59.280 of defining that connectivity.

403 00:19:59.280 --> 00:20:01.890 What it meant by connectivity now basically means

404 00:20:01.890 --> 00:20:05.670 that this function omega ij, if it's non zero,

405 00:20:05.670 --> 00:20:06.780 then that means that there's an edge

406 00:20:06.780 --> 00:20:09.297 between neuron j and neuron I.

407 00:20:09.297 --> 00:20:11.280 And that's basically what I was showing you

408 00:20:11.280 --> 00:20:13.230 in that bigger module.

409 00:20:13.230 --> 00:20:14.640 It all comes down to estimating

410 00:20:14.640 --> 00:20:19.617 whether omega ij is zero or not for this Hawkes process.

411 00:20:20.600 --> 00:20:21.433 Okay.

412 00:20:22.530 --> 00:20:24.810 Let me show you a zero simple example

413 00:20:24.810 --> 00:20:25.650 with two neurons.

414 00:20:25.650 --> 00:20:30.650 In this case, neuron one has no other influence.

415 00:20:32.250 --> 00:20:36.180 It's only it's past history and baseline intensity.

416 00:20:36.180 --> 00:20:40.140 Neuron two has an edge on neuron one.

417 00:20:40.140 --> 00:20:43.430 Let's see what we would expect for the intensity

418 00:20:43.430 --> 00:20:44.280 of neuron one.

419 00:20:44.280 --> 00:20:46.800 If we think about neuro one,

420 00:20:46.800 --> 00:20:50.550 then it's basically a baseline intensity, that new one.

421 00:20:50.550 --> 00:20:55.550 And it's gonna fire at random times for some process.

422 00:20:56.040 --> 00:20:59.481 It's gonna fire at random times with the same intensity.

423 00:20:59.481 --> 00:21:02.040 The intensity is not gonna change because fixed,

424 00:21:02.040 --> 00:21:05.070 we could allow that intensity to be time varying, et cetera,

425 00:21:05.070 --> 00:21:08.130 make it more complicated but in it simplest form

426 00:21:08.130 --> 00:21:11.010 that neuron is just gonna fire randomly,

427 00:21:11.010 --> 00:21:14.103 every time that they sort of it wants.

428 00:21:15.180 --> 00:21:18.600 Now, neuron two would have a difference story

429 00:21:18.600 --> 00:21:22.440 because neuron two depends on activation of neuro one.

430 00:21:22.440 --> 00:21:27.440 Any time that neural one fires, the intensity of neuron two

431 00:21:27.810 --> 00:21:31.230 goes from, let's say the baseline is zero for neuron two,

432 00:21:31.230 --> 00:21:32.760 but every time that neuron one fires,

433 00:21:32.760 --> 00:21:35.700 the intensity of neuron two becomes non zero

434 00:21:35.700 --> 00:21:38.310 because it got excitement from neuron one.

435 00:21:38.310 --> 00:21:39.797 It responds to that.

436 00:21:39.797 --> 00:21:42.330 Neuron two would require to, and then when you have

437 00:21:42.330 --> 00:21:44.880 like three activations, you can get

438 00:21:44.880 --> 00:21:48.480 the convolution of effects that would make neuron two

439 00:21:48.480 --> 00:21:53.480 more likely to activate as well or to spike as well.

440 00:21:53.880 --> 00:21:56.310 And then so this is a pattern that sort of basically

441 00:21:56.310 --> 00:21:58.290 what we are doing here is that we're taking

442 00:21:58.290 --> 00:21:59.680 this to be on omega

443 00:22:01.650 --> 00:22:05.310 to one, that sort of this you see there's the K form

444 00:22:05.310 --> 00:22:08.760 and these get involved if you have more activation

445 00:22:08.760 --> 00:22:11.910 on neuron one, that sort of increases the intensity

446 00:22:11.910 --> 00:22:15.630 of neuron two, meaning that we have more of a chance

447 00:22:15.630 --> 00:22:17.230 for neuron two to fire and this.

448 00:22:20.152 --> 00:22:22.890 Say this simple example, this could be the intensity

449 00:22:22.890 --> 00:22:24.390 of neuron two.

450 00:22:24.390 --> 00:22:28.950 And in fact this all we observe in this case

451 00:22:28.950 --> 00:22:31.670 are these two spike trains for neuron one and neuron two.

452 00:22:31.670 --> 00:22:33.183 We don't observe the network,

453 00:22:34.890 --> 00:22:36.990 in this case there are four possible edges.

454 00:22:36.990 --> 00:22:38.220 One of them is the right edge.

455 00:22:38.220 --> 00:22:41.040 We don't observe the intensity processes.

456 00:22:41.040 --> 00:22:45.420 All we observe is just the point process, the spike.

457 00:22:45.420 --> 00:22:47.460 And the goal is to estimate the network

458 00:22:47.460 --> 00:22:49.440 based on that spike train.

459 00:22:49.440 --> 00:22:50.273 And in fact,

460 00:22:52.980 --> 00:22:56.463 as part of that, we also need to estimate that process.

461 00:23:01.410 --> 00:23:04.593 That estimation problem is not actually that complicated.

462 00:23:05.580 --> 00:23:08.620 If you think of it, it's trying to predict

463 00:23:09.990 --> 00:23:11.433 now based on past.

464 00:23:12.630 --> 00:23:13.680 We could do prediction.

465 00:23:13.680 --> 00:23:17.779 We could use basically penalized regression.

466 00:23:17.779 --> 00:23:19.680 It's a penalized Poison regression.

467 00:23:19.680 --> 00:23:20.820 Something along those lines.

468 00:23:20.820 --> 00:23:21.720 A little bit more complicated,

469 00:23:21.720 --> 00:23:23.697 but basically it's a penalized Poisson regression

470 00:23:23.697 --> 00:23:26.550 and we could use the approach similar

471 00:23:26.550 --> 00:23:28.260 to what is known as neighborhood selection.

472 00:23:28.260 --> 00:23:31.050 We basically meaning that we regress each neuron

473 00:23:31.050 --> 00:23:32.610 on the past of all other neurons,

474 00:23:32.610 --> 00:23:34.290 including that neuron itself.

475 00:23:34.290 --> 00:23:36.331 It's a simple regression problems.

476 00:23:36.331 --> 00:23:39.210 And then we use regularization to select a subset of them

477 00:23:39.210 --> 00:23:42.300 that are more informative, et cetera.

478 00:23:42.300 --> 00:23:44.550 And there's been quite a bit of work on this,

479 00:23:44.550 --> 00:23:46.920 including some work that we've done.

480 00:23:46.920 --> 00:23:49.380 The work that we've done was focused more

481 00:23:49.380 --> 00:23:54.380 on extending the theory of these Hawkes processes

482 00:23:55.100 --> 00:23:57.630 to a setting that is more useful

483 00:23:57.630 --> 00:23:59.820 for neuroscience applications.

484 00:23:59.820 --> 00:24:04.820 In particular, the theory that existed was focused mostly

485 00:24:06.027 --> 00:24:10.530 on the simple linear functions, but also on the case

486 00:24:10.530 --> 00:24:13.770 where we had non-negative transfer functions.

487 00:24:13.770 --> 00:24:17.310 And this was purely an artifact

488 00:24:17.310 --> 00:24:22.200 that the theoretical analysis approach that Hawkes had taken

489 00:24:22.200 --> 00:24:25.413 and using these what are known as cluster representation.

490 00:24:27.690 --> 00:24:32.690 What Hawkes and Oakes had done was that they were

491 00:24:32.910 --> 00:24:37.277 representing each neuron as a sum of, sorry,

492 00:24:39.120 --> 00:24:40.653 homogeneous Poisson processes,

493 00:24:42.303 --> 00:24:44.100 activation pattern of each neuron

494 00:24:44.100 --> 00:24:45.500 as some of homogeneous Poisson process.

495 00:24:45.500 --> 00:24:48.300 And because there was a sum that could not allow

496 00:24:48.300 --> 00:24:51.197 for omega ijs to be negative,

497 00:24:51.197 --> 00:24:55.890 'cause they would cancel throughout and we would get less.

498 00:24:55.890 --> 00:24:59.373 What we did, and this was the work of my former student,

499 00:25:00.330 --> 00:25:03.520 Chen Chang who's Davis, was to

500 00:25:05.820 --> 00:25:08.640 come up with an alternative framework,

501 00:25:08.640 --> 00:25:10.227 theoretical framework motivated by the fact that

502 00:25:10.227 --> 00:25:15.227 we know that neuroscience activations are not just positive,

503 00:25:15.480 --> 00:25:17.550 they're not all excitement,

504 00:25:17.550 --> 00:25:20.133 they're also inhibitions happening.

505 00:25:21.480 --> 00:25:23.790 Neuroscience and in any other biological system really,

506 00:25:23.790 --> 00:25:27.900 we can't have biological systems being stable

507 00:25:27.900 --> 00:25:29.460 without negative feedback.

508 00:25:29.460 --> 00:25:32.370 These negative feedback groups are critical.

509 00:25:32.370 --> 00:25:36.000 We wanted to allow for negative effects

510 00:25:36.000 --> 00:25:37.980 or the effects of inhibition.

511 00:25:37.980 --> 00:25:39.960 And so we came up with a different representation

512 00:25:39.960 --> 00:25:43.530 based on what is known as thinning process representation

513 00:25:43.530 --> 00:25:47.550 that then allowed us to get a concentration

514 00:25:47.550 --> 00:25:48.383 for general.

515 00:25:48.383 --> 00:25:49.590 I won't go into details of this,

516 00:25:49.590 --> 00:25:53.460 that basically we get something that we can show

517 00:25:53.460 --> 00:25:58.460 that for any sort of function,

518 00:25:58.830 --> 00:26:01.443 we get a concentration around its need in a sense.

519 00:26:02.550 --> 00:26:05.730 And so using this as an application,

520 00:26:05.730 --> 00:26:08.250 then you could show that sort of with high probability,

521 00:26:08.250 --> 00:26:10.740 we get to estimate the network correctly

522 00:26:10.740 --> 00:26:14.703 using this name of selection type approach.

523 00:26:15.660 --> 00:26:20.130 This is estimation but we don't really

524 00:26:20.130 --> 00:26:24.350 have any sense of whether...

525 00:26:26.520 --> 00:26:29.190 Let's skip over this for the sake of time.

526 00:26:29.190 --> 00:26:30.870 You don't really have any sense of whether

527 00:26:30.870 --> 00:26:32.850 the edges that we estimate are true edges or not.

528 00:26:32.850 --> 00:26:34.770 We don't have a measure of uncertainty.

529 00:26:34.770 --> 00:26:36.570 We have theory that shows that

530 00:26:36.570 --> 00:26:38.670 sort of the pi should be correct

531 00:26:38.670 --> 00:26:42.930 but we wanna maybe get a sense of uncertainty about this.

532 00:26:42.930 --> 00:26:47.930 And so the work that we've been doing more recently

533 00:26:48.150 --> 00:26:50.490 focused on trying to quantify the uncertainty

534 00:26:50.490 --> 00:26:51.870 of these estimates.

535 00:26:51.870 --> 00:26:54.220 And so there's been a lot of work over the past

536 00:26:55.350 --> 00:26:59.430 almost 10 years on trying to develop inference

537 00:26:59.430 --> 00:27:02.550 for these regularized estimation procedures.

538 00:27:02.550 --> 00:27:03.683 And so we're building on these work,

539 00:27:04.950 --> 00:27:06.300 existing work in particular,

540 00:27:06.300 --> 00:27:09.280 we're building on work on

541 00:27:11.280 --> 00:27:14.280 inferences for vector risk processes.

542 00:27:14.280 --> 00:27:16.180 However, there's some differences

543 00:27:17.340 --> 00:27:22.067 most importantly that vector risk processes capture a fixed

544 00:27:24.030 --> 00:27:27.690 and pre-specified lag, whereas in the Hawkes process case,

545 00:27:27.690 --> 00:27:32.690 we have each basically dependence over the entire history.

546 00:27:33.630 --> 00:27:36.393 We don't have a fixed lag and it's all pre-specified.

547 00:27:37.920 --> 00:27:39.900 And also another difference

548 00:27:39.900 --> 00:27:41.700 is that vector auto-aggressive processes

549 00:27:41.700 --> 00:27:42.533 needs pardoning.

550 00:27:43.560 --> 00:27:44.850 Its' observed over this free time,

551 00:27:44.850 --> 00:27:47.910 whereas the Hawkes process is observed

552 00:27:47.910 --> 00:27:49.505 over a continuous time.

553 00:27:49.505 --> 00:27:50.338 It's a continuous time process

554 00:27:50.338 --> 00:27:52.440 and that that adds a little bit of challenge,

555 00:27:52.440 --> 00:27:56.460 but nonetheless, so we use this de-correlated

556 00:27:56.460 --> 00:27:57.450 score testing work

557 00:27:57.450 --> 00:28:00.930 which is based on the work of Ning and Liu.

558 00:28:00.930 --> 00:28:05.930 And what I'm gonna talk about in the next couple of slides

559 00:28:06.570 --> 00:28:10.740 is an inference framework for these Hawkes processes.

560 00:28:10.740 --> 00:28:13.590 Again, what I showed you before,

561 00:28:13.590 --> 00:28:16.020 the simple form of linear Hawkes process

562 00:28:16.020 --> 00:28:19.080 and motivated by your neuroscience applications,

563 00:28:19.080 --> 00:28:22.200 what we can consider is something quite simple,

564 00:28:22.200 --> 00:28:24.390 although, we could generalize that.

565 00:28:24.390 --> 00:28:26.430 And that generalization is in the paper

566 00:28:26.430 --> 00:28:30.360 but the simple case is to consider something like omega ij

567 00:28:30.360 --> 00:28:34.330 as beta ij times some function pathway j

568 00:28:34.330 --> 00:28:39.330 where that function is simply decay function over time.

569 00:28:40.170 --> 00:28:43.290 It's like exponentially decaying function.

570 00:28:43.290 --> 00:28:44.763 It's class decay function.

571 00:28:45.600 --> 00:28:48.450 That's called a transition for neuroscience applications.

572 00:28:49.290 --> 00:28:52.840 And so if we go with this framework then that

573 00:28:54.224 --> 00:28:57.900 beta ij coefficient determines the connectivity for us,

574 00:28:57.900 --> 00:28:59.853 that this beta ij, if it's positive,

575 00:29:00.750 --> 00:29:03.180 that means that sort of there's an excitement effect.

576 00:29:03.180 --> 00:29:04.857 If it's negative, there's an inhibition effect,

577 00:29:04.857 --> 00:29:08.187 and if it's zero, there's no influence from one or data.

578 00:29:08.187 --> 00:29:11.160 All we need to do really is to develop inference

579 00:29:11.160 --> 00:29:12.153 for this beta ij.

580 00:29:14.340 --> 00:29:17.340 And so that is our goal.

581 00:29:17.340 --> 00:29:22.340 And to do that, I'll go into a little bit of technicalities

582 00:29:22.590 --> 00:29:24.600 and detail of not enough too much.

583 00:29:24.600 --> 00:29:26.880 Please stop me if there are any questions.

584 00:29:26.880 --> 00:29:29.280 The first thing we do is that we realize

585 00:29:29.280 --> 00:29:33.840 that we can represent that linear Hawkes process

586 00:29:33.840 --> 00:29:37.860 as a form of basically a regression almost.

587 00:29:37.860 --> 00:29:41.020 The first thing we do is we turn it into this

588 00:29:43.830 --> 00:29:45.780 integrated stochastic process.

589 00:29:45.780 --> 00:29:47.770 We integrate all the past

590 00:29:48.930 --> 00:29:51.030 that form that sort of seemed ugly,

591 00:29:51.030 --> 00:29:53.400 we integrate it so that it becomes

592 00:29:53.400 --> 00:29:54.780 a little bit more compact.

593 00:29:54.780 --> 00:29:58.500 And then once we do that, we then write it pretty similar

594 00:29:58.500 --> 00:29:59.333 to regression.

595 00:29:59.333 --> 00:30:01.140 We do a change of variable basically.

596 00:30:01.140 --> 00:30:06.140 We write that point process dNi as as our outcome Yi

597 00:30:06.870 --> 00:30:11.100 and then we write epsilon i to be Yi minus lambda

598 00:30:11.100 --> 00:30:14.640 to be added subtract lambda i sense.

599 00:30:14.640 --> 00:30:18.450 And that allows us to write things

600 00:30:18.450 --> 00:30:20.823 as a simple form of regression.

601 00:30:21.810 --> 00:30:24.008 Now this is something that's easy

602 00:30:24.008 --> 00:30:25.470 and we're able to deal with.

603 00:30:25.470 --> 00:30:28.350 The main complication is that sort of this a regression

604 00:30:28.350 --> 00:30:31.500 with the hetero stochastic noise.

605 00:30:31.500 --> 00:30:36.210 Sigma it squared depends on the past

606 00:30:36.210 --> 00:30:38.280 this also time period.

607 00:30:38.280 --> 00:30:40.513 It depends on the beta lambda.

608 00:30:41.850 --> 00:30:44.290 Okay, so once we do this

609 00:30:48.630 --> 00:30:50.943 then to develop a test for beta ij,

610 00:30:53.160 --> 00:30:54.567 we could develop a test for beta ij

611 00:30:54.567 --> 00:30:59.567 and then this also could extended to testing multiple betas

612 00:30:59.580 --> 00:31:02.550 and sort of allowing for ground expansions et cetera.

613 00:31:02.550 --> 00:31:05.880 And even nonstationary the baseline,

614 00:31:05.880 --> 00:31:08.230 but the test is basically

615 00:31:09.270 --> 00:31:11.100 now based on this de-correlated score test.

616 00:31:11.100 --> 00:31:12.810 Once we write in this regression form,

617 00:31:12.810 --> 00:31:15.120 we can take this de-correlated score test

618 00:31:15.120 --> 00:31:18.750 and I'll skip over the details here

619 00:31:18.750 --> 00:31:23.280 but basically we form this set of octagonal columns

620 00:31:23.280 --> 00:31:26.310 and define a score test based on this

621 00:31:26.310 --> 00:31:27.750 that looks something like this,

622 00:31:27.750 --> 00:31:32.163 that you're looking at the effect of the correlated j

623 00:31:32.163 --> 00:31:35.670 with basically noise term, epsilon i.

624 00:31:35.670 --> 00:31:40.200 Both of these are driven from data based on some parameters,

625 00:31:40.200 --> 00:31:42.660 but once you have this, this Sij

626 00:31:42.660 --> 00:31:45.340 then you could actually now define a test

627 00:31:46.770 --> 00:31:51.770 that basically looks at the magnitude of that Sij.

628 00:31:53.340 --> 00:31:56.373 And that's the support that we could use.

629 00:31:59.133 --> 00:32:01.570 And under the no, we can show that this test SUT

630 00:32:01.570 --> 00:32:04.120 converges to a pi square distribution

631 00:32:05.444 --> 00:32:07.530 and we could use that for testing.

632 00:32:07.530 --> 00:32:10.350 In practice, you need to estimate these parameters.

633 00:32:10.350 --> 00:32:12.810 We estimate them, we ensure that things still work

634 00:32:12.810 --> 00:32:14.790 with the estimated parameters

635 00:32:14.790 --> 00:32:17.883 and still so that you have can register pi squared.

636 00:32:19.380 --> 00:32:22.713 And you can also do confidence and all this sector.

637 00:32:23.920 --> 00:32:25.650 Maybe I'll just briefly mention

638 00:32:25.650 --> 00:32:28.980 that this also has the usual power that we expect

639 00:32:28.980 --> 00:32:33.980 that you can study power of this as a local alternative.

640 00:32:34.710 --> 00:32:39.710 And this gives us basically how that we would expect.

641 00:32:41.370 --> 00:32:44.730 And simulation also behaves very close

642 00:32:44.730 --> 00:32:47.460 to the oracle procedure that knows which neurons

643 00:32:47.460 --> 00:32:48.360 acting with other.

644 00:32:49.710 --> 00:32:50.970 What we've done here is that

645 00:32:50.970 --> 00:32:54.270 we've looked at increasing sample size

646 00:32:54.270 --> 00:32:57.597 or own length of the sequence from 200 to 2,000

647 00:32:57.597 --> 00:33:00.690 and then we see that sort of type one error

648 00:33:00.690 --> 00:33:04.710 becomes pretty well controlled as time increases.

649 00:33:04.710 --> 00:33:06.300 The pink here is oracle.

650 00:33:06.300 --> 00:33:07.620 The blue is our procedure.

651 00:33:07.620 --> 00:33:12.620 The power also increases as the sample size increases.

652 00:33:13.560 --> 00:33:17.640 And also look at the coverage of the confidence involved.

653 00:33:17.640 --> 00:33:20.790 Both for the zeros and non zeros,

654 00:33:20.790 --> 00:33:24.033 the coverage also seems to be well behaved.

655 00:33:26.430 --> 00:33:30.700 This is simple setting of simulation but that looks like

656 00:33:32.010 --> 00:33:35.340 it's not too far actually in application

657 00:33:35.340 --> 00:33:36.640 that we've also looked at.

658 00:33:38.027 --> 00:33:40.900 And in particular we've looked at some data

659 00:33:41.940 --> 00:33:44.880 paper that was published in 2018 in nature

660 00:33:44.880 --> 00:33:49.880 when they had looked at activation patterns of neurons

661 00:33:50.070 --> 00:33:52.923 and how they would change with and without laser.

662 00:33:54.002 --> 00:33:56.640 And at the time this was like the largest,

663 00:33:56.640 --> 00:33:59.547 so they had multiple device that they had looked at,

664 00:33:59.547 --> 00:34:01.860 and this was the largest region

665 00:34:01.860 --> 00:34:04.320 that they had looked at had 25 neurons.

666 00:34:04.320 --> 00:34:05.760 The technology has improved quite a bit.

667 00:34:05.760 --> 00:34:07.500 Now there's a couple of hundred neurons

668 00:34:07.500 --> 00:34:09.300 that they could measure,

669 00:34:09.300 --> 00:34:10.133 but this was 25 neurons.

670 00:34:10.133 --> 00:34:13.530 And then what I'm showing you are the activation patterns

671 00:34:13.530 --> 00:34:15.810 without laser and with laser

672 00:34:15.810 --> 00:34:18.900 and not showing the edges that are common

673 00:34:18.900 --> 00:34:19.980 between the two networks.

674 00:34:19.980 --> 00:34:21.120 I'm just showing the edges are different

675 00:34:21.120 --> 00:34:22.810 between these networks.

676 00:34:22.810 --> 00:34:25.290 And we see that these betas,

677 00:34:25.290 --> 00:34:27.540 some of them are clearly different.

678 00:34:27.540 --> 00:34:31.530 In one condition the coefficient covers zero

679 00:34:31.530 --> 00:34:32.850 and the other conditions not cover.

680 00:34:32.850 --> 00:34:35.547 And that's why you're seeing these difference in networks.

681 00:34:35.547 --> 00:34:38.550 And that's similar to what they had observed

682 00:34:38.550 --> 00:34:43.440 based on basically correlation that as you activate

683 00:34:43.440 --> 00:34:46.173 there's more connectivity among these neurons.

684 00:34:48.540 --> 00:34:51.300 Now in the actual experiments,

685 00:34:51.300 --> 00:34:56.300 and this is maybe the last 15 minutes or so by top,

686 00:34:57.300 --> 00:35:00.090 in the actual experiments, they don't do just a simple

687 00:35:00.090 --> 00:35:02.610 one shot experiment because they have to implant

688 00:35:02.610 --> 00:35:03.663 this device.

689 00:35:06.030 --> 00:35:07.830 This is data of a mouse.

690 00:35:07.830 --> 00:35:10.980 They have to implant this device on mouse's brain.

691 00:35:10.980 --> 00:35:12.810 And so what they do is that they actually,

692 00:35:12.810 --> 00:35:16.320 once they do that and sort of now with that camera,

693 00:35:16.320 --> 00:35:18.330 they just measure activities of neurons.

694 00:35:18.330 --> 00:35:20.370 But once they do that, they actually run

695 00:35:20.370 --> 00:35:22.530 a sequence of experiments.

696 00:35:22.530 --> 00:35:25.170 It's never just a single experiment or two experiments.

697 00:35:25.170 --> 00:35:28.170 What they do is that they, for example,

698 00:35:28.170 --> 00:35:31.140 they show different images, the mouse

699 00:35:31.140 --> 00:35:34.050 and they see the activation patterns of neurons

700 00:35:34.050 --> 00:35:36.090 as the mouse processes different images.

701 00:35:36.090 --> 00:35:37.950 And what they usually do is that sort they show an image

702 00:35:37.950 --> 00:35:41.940 with one orientation and then they have a washout period.

703 00:35:41.940 --> 00:35:43.743 They show an image with different orientation,

704 00:35:43.743 --> 00:35:44.723 they have a washout period.

705 00:35:44.723 --> 00:35:46.620 They show an image with a different orientation

706 00:35:46.620 --> 00:35:49.680 and then they might use laser

707 00:35:49.680 --> 00:35:52.803 in combination of these different images et cetera.

708 00:35:52.803 --> 00:35:54.060 What they ended up doing

709 00:35:54.060 --> 00:35:56.220 is that they have many, many experiments.

710 00:35:56.220 --> 00:35:58.680 And what we expect is that the networks

711 00:35:58.680 --> 00:35:59.780 in these different experiments

712 00:35:59.780 --> 00:36:01.500 to be different from each other

713 00:36:01.500 --> 00:36:04.470 but maybe share some commonalities as well.

714 00:36:04.470 --> 00:36:06.240 We don't expect completely different networks

715 00:36:06.240 --> 00:36:08.343 but we expect somewhat related networks.

716 00:36:09.270 --> 00:36:13.470 And over different time segments

717 00:36:13.470 --> 00:36:14.880 the network might change.

718 00:36:14.880 --> 00:36:18.510 In one segment it might be that and the next segment

719 00:36:18.510 --> 00:36:20.250 it might change to something different

720 00:36:20.250 --> 00:36:23.073 but maybe some parts of the network structure are like.

721 00:36:24.660 --> 00:36:26.670 What this does is that it sort of motivates us

722 00:36:26.670 --> 00:36:28.860 to think about join the estimate in these networks

723 00:36:28.860 --> 00:36:31.110 because each one of these time segments

724 00:36:31.110 --> 00:36:34.890 might not have enough observation to estimate accurately.

725 00:36:34.890 --> 00:36:36.227 And this goes back to the simulation results

726 00:36:36.227 --> 00:36:40.710 that I showed you, that in order to get to good control

727 00:36:40.710 --> 00:36:42.720 of type one error and good power,

728 00:36:42.720 --> 00:36:44.670 we need to have decent number of observations.

729 00:36:44.670 --> 00:36:46.920 And in each one of these time segments

730 00:36:46.920 --> 00:36:48.813 might not have enough observations.

731 00:36:50.460 --> 00:36:54.270 In order to make sure that we get high quality estimates

732 00:36:54.270 --> 00:36:57.180 and valid inference,

733 00:36:57.180 --> 00:36:59.730 we need to maybe join the estimations

734 00:36:59.730 --> 00:37:04.173 in order to get better quality estimates and influence.

735 00:37:11.130 --> 00:37:13.392 That's the idea of the second part

736 00:37:13.392 --> 00:37:16.950 of what I wanna talk about going beyond

737 00:37:16.950 --> 00:37:19.290 the single experiment and trying to do estimation

738 00:37:19.290 --> 00:37:22.380 and inference, and multiple experiments of similar.

739 00:37:22.380 --> 00:37:26.010 And in fact in the case of this paper by and Franks

740 00:37:26.010 --> 00:37:30.210 they had, for every single mouse,

741 00:37:30.210 --> 00:37:33.300 they had 80 different experimental setups

742 00:37:33.300 --> 00:37:34.830 with laser and different durations

743 00:37:34.830 --> 00:37:36.540 and different strengths.

744 00:37:36.540 --> 00:37:39.210 It's not a single experiment for each mouse.

745 00:37:39.210 --> 00:37:41.610 It's 80 different experiments for each mouse.

746 00:37:41.610 --> 00:37:44.190 And you would expect that many of these experiments

747 00:37:44.190 --> 00:37:45.300 are similar to each other

748 00:37:45.300 --> 00:37:47.280 and they might have different degrees of similarities

749 00:37:47.280 --> 00:37:50.317 with each other that might need to take into account.

750 00:37:52.713 --> 00:37:55.740 Then the goal of the second part is do joint estimation

751 00:37:55.740 --> 00:37:59.040 of inference for settings where we have multiple experiments

752 00:37:59.040 --> 00:38:00.690 and not just a single experiment.

753 00:38:01.800 --> 00:38:04.620 To do this, we went back to basically

754 00:38:04.620 --> 00:38:06.570 that destination that we had

755 00:38:06.570 --> 00:38:10.530 and previously what we had was the sparsity type penalty.

756 00:38:10.530 --> 00:38:12.150 What we do is that sort of now we added

757 00:38:12.150 --> 00:38:13.560 a fusion type penalty.

758 00:38:13.560 --> 00:38:17.323 Now we combine the estimates in different experiments.

759 00:38:18.840 --> 00:38:22.200 And this is based on past work that I had done

760 00:38:22.200 --> 00:38:23.730 with the the post

761 00:38:23.730 --> 00:38:26.470 but the main difference in this board is that

762 00:38:27.840 --> 00:38:31.620 now we wanna allow these estimates

763 00:38:31.620 --> 00:38:33.420 to be similar to each other

764 00:38:33.420 --> 00:38:35.760 based on a data-driven notion of similarity.

765 00:38:35.760 --> 00:38:37.050 We don't know which experiments

766 00:38:37.050 --> 00:38:39.677 are more similar to each other.

767 00:38:39.677 --> 00:38:43.320 And we basically want the data to tell us which experiments

768 00:38:43.320 --> 00:38:45.720 should be more similar to each other, should be combined

769 00:38:45.720 --> 00:38:50.720 and not necessarily find that a priority person

770 00:38:50.820 --> 00:38:52.719 usually don't have that information.

771 00:38:52.719 --> 00:38:57.120 These data-driven weights are critical here,

772 00:38:57.120 --> 00:38:59.190 and we drive these data-driven weights

773 00:38:59.190 --> 00:39:00.960 based on just simple correlations.

774 00:39:00.960 --> 00:39:02.160 We calculate simple correlations.

775 00:39:02.160 --> 00:39:05.370 The first step we look to see which one of these conditions,

776 00:39:05.370 --> 00:39:08.575 the correlations are more correlated with each other,

777 00:39:08.575 --> 00:39:10.680 more similar to each other

778 00:39:10.680 --> 00:39:12.570 based on these correlations.

779 00:39:12.570 --> 00:39:17.190 And we use these cost correlations to then define ways

780 00:39:17.190 --> 00:39:19.650 for which experiments should be more closely used

781 00:39:19.650 --> 00:39:20.580 with each other.

782 00:39:20.580 --> 00:39:22.050 And estimates on which experiments

783 00:39:22.050 --> 00:39:24.540 should be more closely used.

784 00:39:24.540 --> 00:39:28.770 And I leave that in terms of details

785 00:39:28.770 --> 00:39:32.400 but in this similar setting

786 00:39:32.400 --> 00:39:34.320 as what I had explained before

787 00:39:34.320 --> 00:39:36.870 in terms of experimental setup for this,

788 00:39:36.870 --> 00:39:39.210 I'm sorry, in terms of simulation setup,

789 00:39:39.210 --> 00:39:41.703 there are 50 neurons in network

790 00:39:41.703 --> 00:39:44.040 from three different experiments in this case

791 00:39:44.040 --> 00:39:45.450 of three different lengths,

792 00:39:45.450 --> 00:39:47.820 and we use different estimators.

793 00:39:47.820 --> 00:39:51.060 And what we see is that sort of when we do this fusion,

794 00:39:51.060 --> 00:39:54.480 we do better in terms of the number of two positives

795 00:39:54.480 --> 00:39:57.090 for any given number of estimated edges

796 00:39:57.090 --> 00:39:59.250 compared to separately estimating

797 00:39:59.250 --> 00:40:02.430 or compared to sort of other types of fusions

798 00:40:02.430 --> 00:40:04.113 that what one might consider.

799 00:40:05.940 --> 00:40:10.110 Now, estimation is somewhat easy.

800 00:40:10.110 --> 00:40:11.610 The main challenge was to come up

801 00:40:11.610 --> 00:40:13.980 with these data-driven weights.

802 00:40:13.980 --> 00:40:17.830 The main issue is that if you wanted to come up with

803 00:40:19.290 --> 00:40:20.850 valid infants in these settings,

804 00:40:20.850 --> 00:40:24.330 when we have many, many experiments,

805 00:40:24.330 --> 00:40:26.670 then then we would have very low power if we're adjusting,

806 00:40:26.670 --> 00:40:29.777 for example, from all comparison using FDR, FWER,

807 00:40:31.261 --> 00:40:33.783 false discovery rate or family-wise error rate,

808 00:40:35.010 --> 00:40:37.380 we have p squared times MS.

809 00:40:37.380 --> 00:40:39.840 And so we have a low power.

810 00:40:39.840 --> 00:40:41.790 To deal with this setting, what we have done

811 00:40:41.790 --> 00:40:45.180 is that we've come up with a hierarchical testing procedure

812 00:40:45.180 --> 00:40:48.970 that avoids testing

813 00:40:49.890 --> 00:40:52.285 all these p squared times M coefficient.

814 00:40:52.285 --> 00:40:53.118 And the idea is this,

815 00:40:53.118 --> 00:40:56.580 the idea is that if you have a sense of which conditions

816 00:40:56.580 --> 00:40:58.560 are more similar to each other,

817 00:40:58.560 --> 00:41:03.000 we construct a very specific type of binary tree,

818 00:41:03.000 --> 00:41:06.660 which basically always has a single node

819 00:41:06.660 --> 00:41:09.092 on the left side in this case.

820 00:41:09.092 --> 00:41:10.767 And then we start on the top of that tree

821 00:41:10.767 --> 00:41:13.050 and and test for each coefficient.

822 00:41:13.050 --> 00:41:15.620 We first test Albany experiments.

823 00:41:15.620 --> 00:41:18.330 If you don't reject, then you stop there.

824 00:41:18.330 --> 00:41:22.260 If you reject then we test one, and two,

825 00:41:22.260 --> 00:41:24.720 three, and four separately.

826 00:41:24.720 --> 00:41:28.080 If you reject one, then we've identified the non

827 00:41:28.080 --> 00:41:30.150 make the non zero edge.

828 00:41:30.150 --> 00:41:33.817 If you reject two, three, four, then we go down.

829 00:41:33.817 --> 00:41:36.060 If you don't reject two, three, four, we stop there.

830 00:41:36.060 --> 00:41:39.270 This way we stop at the level that is appropriate

831 00:41:39.270 --> 00:41:40.263 based on data.

832 00:41:42.193 --> 00:41:44.370 And this this ends up especially in sparse networks,

833 00:41:44.370 --> 00:41:47.530 this ends up saving us a lot of tests

834 00:41:48.838 --> 00:41:51.150 and gives us significant improvement in power.

835 00:41:51.150 --> 00:41:53.370 And that's shown in the simulation

836 00:41:53.370 --> 00:41:57.000 that you end up, if you don't do this,

837 00:41:57.000 --> 00:42:00.570 your power decreases as the number of experiments increases.

838 00:42:00.570 --> 00:42:03.660 And in this case you've gone up to 50 experiments

839 00:42:03.660 --> 00:42:04.493 as I mentioned.

840 00:42:04.493 --> 00:42:07.140 The golden and facts paper has about 80.

841 00:42:07.140 --> 00:42:08.637 Whereas if you don't do that

842 00:42:08.637 --> 00:42:10.983 and if your network sparse actually power,

843 00:42:12.330 --> 00:42:14.970 you see that by combining experiments,

844 00:42:14.970 --> 00:42:15.900 you actually gain power

845 00:42:15.900 --> 00:42:17.850 because you're incorporating more data.

846 00:42:18.870 --> 00:42:22.096 And this is more controlling the family-wise error rate.

847 00:42:22.096 --> 00:42:25.020 And both methods control the famil-wise error rate.

848 00:42:25.020 --> 00:42:26.790 We haven't developed anything for FDR.

849 00:42:26.790 --> 00:42:28.950 We haven't developed theory for FDR

850 00:42:28.950 --> 00:42:31.582 but the method also seems to be controlling FDR

851 00:42:31.582 --> 00:42:34.916 in a very stringent way actually.

852 00:42:34.916 --> 00:42:38.130 But we just don't have theory for FDR control

853 00:42:38.130 --> 00:42:39.980 'cause that becomes more complicated.

854 00:42:45.930 --> 00:42:47.430 I'm going very fast because of time

855 00:42:47.430 --> 00:42:49.410 but I'll pause for a minute.

856 00:42:49.410 --> 00:42:50.243 Any questions.

857 00:42:53.010 --> 00:42:54.240 Please.

858 00:42:54.240 --> 00:42:56.400 <v ->What do you think about stationary</v>

859 00:42:56.400 --> 00:42:58.110 of the Hawkes process in the context?

860 00:42:58.110 --> 00:43:01.050 Whether it’s the exogenous experimental forcing

861 00:43:01.050 --> 00:43:02.960 and like over what timescale did that happen

862 00:43:02.960 --> 00:43:04.470 in the stationary, the reasonable?

863 00:43:04.470 --> 00:43:06.370 <v ->Yeah, that’s a really good question.</v>

864 00:43:10.845 --> 00:43:12.810 To be honest, I think these hard processes

865 00:43:12.810 --> 00:43:14.490 are most likely non stationary.

866 00:43:14.490 --> 00:43:19.490 The two mechanisms of non stationary that could happen.

867 00:43:19.710 --> 00:43:22.050 One, we try to account for it.

868 00:43:22.050 --> 00:43:24.788 I skipped over it but we tried to account

869 00:43:24.788 --> 00:43:27.750 for one aspect of it by allowing the baseline rate

870 00:43:27.750 --> 00:43:29.793 to be time varying.

871 00:43:37.555 --> 00:43:42.555 Basically we allow this this new i to be a function of time.

872 00:43:42.810 --> 00:43:47.730 Baseline rate for each neuron is varying over time.

873 00:43:47.730 --> 00:43:49.320 And the hope is that, that would capture

874 00:43:49.320 --> 00:43:53.313 some of the exogenous factors that might influence overall.

875 00:43:55.857 --> 00:44:00.150 It could also be that the data are changing over time.

876 00:44:00.150 --> 00:44:04.787 That sort of we haven’t done or it could in fact be that

877 00:44:06.150 --> 00:44:08.710 we have abrupt changes

878 00:44:10.200 --> 00:44:14.637 in patterns of either activation or the baseline over time,

879 00:44:14.637 --> 00:44:16.620 but sort all of a sudden something completely changes.

880 00:44:16.620 --> 00:44:21.620 We have piecewise stationary, not monotone sort of,

881 00:44:22.050 --> 00:44:23.891 not continuous, not stationary.

882 00:44:23.891 --> 00:44:25.890 We have piecewise.

883 00:44:25.890 --> 00:44:27.690 We have experimental that's happening,

884 00:44:27.690 --> 00:44:29.520 something happening and then all of a sudden

885 00:44:29.520 --> 00:44:31.110 something else is happening.

886 00:44:31.110 --> 00:44:35.182 This eventually would capture maybe plasticity

887 00:44:35.182 --> 00:44:38.670 in these neurons to neuroplasticity to some extent

888 00:44:38.670 --> 00:44:42.120 that sort of allows for changes of activity over time,

889 00:44:42.120 --> 00:44:44.103 but beyond that we haven't done any.

890 00:44:45.090 --> 00:44:46.710 There's actually one paper that has looked

891 00:44:46.710 --> 00:44:49.923 at piece stationary for these hard processes neuron.

892 00:44:52.260 --> 00:44:55.010 It becomes a competition, very, very difficult problem,

893 00:44:55.890 --> 00:44:59.105 especially the person becomes very difficult problem.

894 00:44:59.105 --> 00:45:01.005 But I think it's a very good question.

895 00:45:03.030 --> 00:45:06.393 Aside from that one paper much else that has done.

896 00:45:10.980 --> 00:45:12.930 <v ->Hi, thank you professor for the sharing.</v>

897 00:45:12.930 --> 00:45:15.130 I have a question regarding the segmentation

898 00:45:16.827 --> 00:45:19.350 'cause on the video you showed us,

899 00:45:19.350 --> 00:45:22.590 the image is generally very shaky.

900 00:45:22.590 --> 00:45:25.020 In the computer vision perspective,

901 00:45:25.020 --> 00:45:28.260 it's very hard to isolate which neuron actually fired

902 00:45:28.260 --> 00:45:31.590 and make sure that it's that same neuron fires over time.

903 00:45:31.590 --> 00:45:35.940 And also the second question is that the mouse

904 00:45:35.940 --> 00:45:39.060 factory, the model you've mentioned is like 20 neurons,

905 00:45:39.060 --> 00:45:41.520 but in the picture you show us there's probably

906 00:45:41.520 --> 00:45:42.360 thousands of neurons.

907 00:45:42.360 --> 00:45:44.893 How do you identify which 20 neurons to look at?

908 00:45:45.753 --> 00:45:47.850 <v ->Very good questions.</v>

909 00:45:47.850 --> 00:45:50.610 First of all, before they even get to segmentation,

910 00:45:50.610 --> 00:45:52.260 they need to do what is known as,

911 00:45:54.960 --> 00:45:57.820 and this is actually common in

912 00:45:58.950 --> 00:46:00.800 time series and sort of (indistinct).

913 00:46:02.641 --> 00:46:03.974 In registration.

914 00:46:07.071 --> 00:46:09.270 What this means is that you first need to register

915 00:46:09.270 --> 00:46:12.600 the images so that they're basically aligning correct.

916 00:46:12.600 --> 00:46:14.490 Then you can do segmentation.

917 00:46:14.490 --> 00:46:17.310 If you remember first five,

918 00:46:17.310 --> 00:46:19.620 but if you remember had a couple of dots

919 00:46:19.620 --> 00:46:21.000 before getting to segmentation.

920 00:46:21.000 --> 00:46:22.800 There are a couple of steps that need to happen

921 00:46:22.800 --> 00:46:25.050 before we even get to segmentation.

922 00:46:25.050 --> 00:46:26.700 And part of that is registration.

923 00:46:26.700 --> 00:46:28.680 Registration is actually a nontrivial pass

924 00:46:28.680 --> 00:46:31.800 to make sure that the vocations don't change.

925 00:46:31.800 --> 00:46:36.210 You have to right otherwise that the algorithm

926 00:46:36.210 --> 00:46:37.440 will get confused.

927 00:46:37.440 --> 00:46:41.280 First there's a registration that needs to happen

928 00:46:41.280 --> 00:46:42.510 and some background correction

929 00:46:42.510 --> 00:46:45.267 and sort of getting noise correctly and everything.

930 00:46:45.267 --> 00:46:46.680 And then there's registration.

931 00:46:46.680 --> 00:46:48.810 And then after that you could do segmentation,

932 00:46:48.810 --> 00:46:50.040 identifying neurons.

933 00:46:50.040 --> 00:46:52.380 Now, the data that they showed you was a data

934 00:46:52.380 --> 00:46:56.257 from actually cats video that showed it's different,

935 00:46:56.257 --> 00:46:59.727 this holding and banks data that they showed you here.

936 00:46:59.727 --> 00:47:02.550 This one had 25 neurons that they had.

937 00:47:02.550 --> 00:47:04.410 This is an older technology.

938 00:47:04.410 --> 00:47:06.600 It's an older paper that they only had 25 neurons,

939 00:47:06.600 --> 00:47:09.980 that they had smaller regions that they were capturing.

940 00:47:09.980 --> 00:47:11.350 The newer technologies, they were capturing

941 00:47:11.350 --> 00:47:14.130 the larger region a couple hundred.

942 00:47:14.130 --> 00:47:15.578 I think the most I've seen

943 00:47:15.578 --> 00:47:17.310 was about a thousand or so neurons.

944 00:47:17.310 --> 00:47:19.770 I haven't seen more than a thousand neurons.

945 00:47:19.770 --> 00:47:20.603 <v ->Thank you.</v>

946 00:47:25.372 --> 00:47:28.776 <v ->Okay, so I'm close to the end of my time.</v>

947 00:47:28.776 --> 00:47:33.776 Maybe I'll have the remaining minutes or so

948 00:47:34.320 --> 00:47:36.570 I'll basically mention that sort of

949 00:47:36.570 --> 00:47:39.220 give by this saying we have joint estimation

950 00:47:41.820 --> 00:47:42.660 to the data from holding advance.

951 00:47:42.660 --> 00:47:47.610 And then we also see that something that is not surprising

952 00:47:47.610 --> 00:47:50.686 perhaps that the no laser condition,

953 00:47:50.686 --> 00:47:52.838 the net yield is more different

954 00:47:52.838 --> 00:47:55.170 than the two different magnitudes of laser,

955 00:47:55.170 --> 00:48:00.043 maybe 10, 20 sort of meters and so square.

956 00:48:02.100 --> 00:48:04.740 You see that so least two are more similar other

957 00:48:04.740 --> 00:48:07.563 than the no laser condition.

958 00:48:09.791 --> 00:48:11.670 And I'm probably gonna stop here

959 00:48:11.670 --> 00:48:14.010 and sort of leave a couple of minutes for questions,

960 00:48:14.010 --> 00:48:15.300 additional questions, but I'll mention that

961 00:48:15.300 --> 00:48:18.720 so the last part I didn't talk about was to see if we could

962 00:48:18.720 --> 00:48:20.372 go beyond prediction.

963 00:48:20.372 --> 00:48:23.010 Could we use this and mention that sort major causality

964 00:48:23.010 --> 00:48:26.510 is not really causality prediction.

965 00:48:26.510 --> 00:48:29.013 It could we go beyond prediction,

966 00:48:30.930 --> 00:48:34.800 get a sense of which neurons are impacting other neurons.

967 00:48:34.800 --> 00:48:38.850 And I'll briefly mention that sort of there are two issues

968 00:48:38.850 --> 00:48:42.573 in general going beyond prediction causality.

969 00:48:44.640 --> 00:48:47.160 We have a review paper that tlaks about this one,

970 00:48:47.160 --> 00:48:48.348 issue is subsampling.

971 00:48:48.348 --> 00:48:51.300 And that you don't have enough resolution.

972 00:48:51.300 --> 00:48:52.683 And the other issue is where you might have

973 00:48:52.683 --> 00:48:55.470 limited processes that make it difficult

974 00:48:55.470 --> 00:48:57.377 to answer all the questions.

975 00:48:57.377 --> 00:49:00.180 Fortunately the issue of self sampling,

976 00:49:00.180 --> 00:49:04.170 which is a difficult issue in general is not present,

977 00:49:04.170 --> 00:49:07.983 but is not very prominent thinking these classroom

978 00:49:09.269 --> 00:49:10.470 and imaging data

979 00:49:10.470 --> 00:49:14.327 because you have continuous time videos.

980 00:49:14.327 --> 00:49:19.260 And subsampling should not be a big deal in this case.

981 00:49:19.260 --> 00:49:22.530 However, we observe a tiny faction

982 00:49:22.530 --> 00:49:25.290 of the connection of the brain.

983 00:49:25.290 --> 00:49:27.480 The question is, can we somehow account

984 00:49:27.480 --> 00:49:29.680 for all the other neurons that we don't see?

985 00:49:31.260 --> 00:49:34.080 The last part of this work is about that.

986 00:49:34.080 --> 00:49:37.770 And I'll sort of jump to the end

987 00:49:37.770 --> 00:49:40.800 because I'll put a reference to that work.

988 00:49:40.800 --> 00:49:43.020 That one is published in case you're interested

989 00:49:43.020 --> 00:49:46.150 in a paper that sort of looks at

990 00:49:48.855 --> 00:49:50.910 whether we could go beyond prediction,

991 00:49:50.910 --> 00:49:53.760 whether they actually identify causal links

992 00:49:53.760 --> 00:49:54.810 particularly neurons.

993 00:49:55.692 --> 00:49:59.580 And I think I'm gonna stop here and thank you guys

994 00:49:59.580 --> 00:50:01.823 and I'm happy to take more questions.

995 00:50:16.900 --> 00:50:18.063 <v ->Naive question.</v>

996 00:50:19.396 --> 00:50:24.396 Biologically, what is a network connection here?

997 00:50:24.431 --> 00:50:27.150 Because they're not, I'm assuming they're not

998 00:50:27.150 --> 00:50:30.143 growing synapses or not based on the laser.

999 00:50:33.099 --> 00:50:36.271 (indistinct)

1000 00:50:36.271 --> 00:50:39.188 (group chattering)