Decoding and perturbing decision states in real time

2 *Diogo Peixoto[1,2]*, Jessica R. Verhein[3,4]*, Roozbeh Kiani[5], Jonathan C. Kao[6,7,8], Paul*

3 *Nuyujukian[3,6,9,10,15,16], Chandramouli Chandrasekaran[6,11,12,13], Julian Brown[11], Sania Fong[11],*

4 *Stephen I. Ryu[6,14], Krishna V. Shenoy[1,6,9,11,15,16], William T. Newsome[1,11,15,16]*

5 1 Neurobiology Department, Stanford University, Stanford, CA 94305

6 2 Champalimaud Neuroscience Programme, Lisbon, Portugal, 1400-038

7 3 Neurosciences Graduate Program, Stanford University, Stanford, CA 94305

8 4 Medical Scientist Training Program, Stanford University School of Medicine, Stanford, CA

9 94305

10 5 Center for Neural Science, New York University, New York, NY 10003

11 6 Electrical Engineering Department, Stanford University, Stanford, CA 94305

12 7 Department of Electrical and Computer Engineering, University of California, Los Angeles, CA

13 90095

14 8 Neurosciences Program, University of California, Los Angeles, CA 90095

15 9 Bioengineering Department, Stanford University, Stanford, CA 94305

16 10 Neurosurgery Department, Stanford University, Stanford, CA 94305

17 11 Howard Hughes Medical Institute, Stanford University, CA 94305

18 12 Department of Psychological and Brain Sciences, Boston University, Boston, MA 02215

19 13 Department of Anatomy and Neurobiology, Boston University School of Medicine, Boston,

20 MA 02118

21 14 Neurosurgery Department, Palo Alto Medical Foundation, Palo Alto, CA 94301

22 15 Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA 94305

23   16 Bio-X Institute, Stanford University, Stanford, CA 94305

24   * Denotes equal contribution.

25   Please address correspondence to Bill Newsome (bnewsome@stanford.edu), Diogo Peixoto

26   (dpeixoto@stanford.edu) or Jessica Verhein (jverhein@stanford.edu).

27

Summary

29   In dynamic environments, subjects often integrate multiple samples of a signal and combine them

30   to reach a categorical judgment. The process of deliberation on the evidence can be described by

31   a time-varying decision variable (DV), decoded from neural activity, that predicts a subject's

32   decision at the end of a trial. However, within trials, large moment-to-moment fluctuations of the

33   DV are observed. The behavioral significance of these fluctuations and their role in the decision

34   process remain unclear. Here we show that within-trial DV fluctuations decoded in real time from

35   motor cortex are tightly linked to choice behavior, and that robust changes in DV sign have the

36   statistical regularities expected from behavioral studies of changes-of-mind. Furthermore, we find

37   single-trial evidence for absorbing decision bounds. As the DV builds up, heavily favoring one or

38   the other choice, moment-to-moment variability in the DV is reduced, and both neural DV and

39   behavioral decisions become resistant to additional pulses of sensory evidence as predicted by

40   diffusion-to-bound and attractor models of the decision process.

41

42

43

44   When making a categorical decision about a noisy stimulus, it is common to fluctuate between

45   levels of commitment to a choice before reporting a decision. In some instances the fluctuations

46   are sufficiently strong to lead to a "change of mind" (CoM) while deliberating[1-6] or even while the

47   reporting action is being executed[7]. Because these within-trial fluctuations are different from trial

48   to trial and not necessarily tied to an external event or stimulus feature, they can only be captured

49   using a moment-to-moment neural readout of the decision state on single trials.

50   To obtain this readout, we decoded a decision variable (DV) from neural population activity in

51   PMd and M1 in real time to continuously estimate the decision state while two monkeys performed

52   a motion discrimination task[8,9] (Fig. 1a, see Methods). The DV was estimated by applying a linear

53   decoder, trained on data from a previous experimental session, to spiking data (from 96 to 192

54   electrodes) from the preceding 50 ms, updated every 10 ms throughout each trial (Fig. 1b, see

55   Methods). The sign of the DV indicated which choice was predicted by the decoder, which allowed

56   us to calculate the decoder's prediction accuracy. The DV magnitude reflected the confidence of

57   the model's prediction in units of log-odds for one vs. the other decision (see Methods). Note that

58   the decision variable as defined here encompasses all choice predictive signals that can be decoded

59   from neural activity[10], including but not limited to moment-to-moment value of accumulated

60   evidence as posited in classical sequential sampling models.

61   We have previously demonstrated with offline analysis that this decision variable (DV) can predict

62   choices on single trials up to seconds before initiation of the operant response, and that the

63   accuracy of these predictions increases on average throughout the course of the trial[10].

64   Here, we employed closed-loop, neurally-contingent control over stimulus timing to directly probe

65   the relationship of within-trial DV fluctuations to behaviorally meaningful decision states. For the

66   first time, we quantified the behavioral effects of previously covert DV variations (i) as a function

67   of time and for different virtual DV boundaries imposed during the trial, (ii) when large, CoM-like

68   fluctuations were detected during deliberation on noisy visual evidence, and (iii) when

69   subthreshold stimulus pulses were added during the trial.

70   Having a nearly instantaneous real-time estimate of the decision state read-out enabled us to

71   terminate the visual stimulus based on the current value (or history) of the DV and validate the

72   behavioral relevance of DV fluctuations using the monkey's behavioral reports following stimulus

73   termination.

74   **Decisions on perceived stimulus motion can be reliably decoded in real time based on 50 ms**

75   **of PMd/M1 neural activity**

76

77   Two monkeys performed a variable duration variant of the classical random dot motion

78   discrimination task using an arm movement as the operant response[10]. As expected, the subjects

79   performed better for higher coherence and longer duration stimuli and reached almost perfect

80   performance for the easiest stimuli (Extended Data Fig. 1).

81   We first measured the accuracy of our real-time decoder in predicting the monkeys' behavioral

82   choices as a function of time during the trial. As in our previous offline results[10], average prediction

83   accuracy started at chance levels during the targets epoch (Fig. 1c, Extended Data Fig. 2a). During

84   the dots presentation average prediction accuracy quickly departed from baseline (174.5 ms ± 18.8

85   and 214.5 ms ± 8.09 ms after dots onset for monkey H and F, respectively), rising monotonically

86   for the rest of the epoch. The rise in prediction accuracy was steep, reaching 99% (98%) correct

87   for the longest stimuli presentations for monkey H (F), respectively. Moreover, for all 4 epochs

88   considered (targets, dots, delay and post-go) the average accuracy difference between our real-

89   time readout and the equivalent one calculated offline (trained using data from the same session)

90   was within a ± 2% range (Extended Data Fig. 3a-d). Thus, our real-time choice decoder reproduces

91   prediction accuracy as reported in previous off-line analyses of decision-related neural activity in

92   both the oculomotor and somatomotor systems[1,10].

93   Our real-time decoder also reproduced the temporal dynamics and coherence dependence of the

94   DV, as reported in previous off-line studies[1,10]. The on-line DV: (i) started around 0 at the time of

95   dots onset, (ii) separated by choice after ∼200 ms, and (iii) rose (or fell) faster for easier trials (Fig.

96   1d, Extended Data Fig. 2b; regression of DV onto coherence significant for both choices, $p<10^{-5}$

97   uncorrected). Prediction accuracy was higher for correct trials compared to error trials (Extended

98   Data Fig. 4) when holding the stimulus coherence constant, as expected from previous studies[11].

99   Finally, our decoding method yielded stable performance across multiple days, justifying

100   combination of data across sessions (Extended Data Fig. 5). This is particularly important when

101   studying rare events such as CoMs, which only happen on a small fraction of the trials and could

102   not be characterized adequately using a single session's data.

103   **Real time DV closely predicts choice likelihood across experimental conditions**

104   The previous results are a proof of concept for a highly reliable, real-time readout of decision state

105   in PMd/M1 using spiking data from ∼100-200 units and aggregate and average metrics (Fig. 1c-

106   d, Extended Data Fig. 2a-b). However, we often observed large fluctuations (over 3 natural log

107   units) in the decision variable on individual trials, even within single behavioral epochs (Fig. 1e).

108   If moment-to-moment fluctuations in DV during single trials (as estimated by our decoder) reflect

109   true fluctuations in the decision state of the animal, we expect larger absolute values of DV to be

5

110   associated with stronger preference for one of the two choices, and hence higher prediction

111   accuracy were a decision to be required at any time during a single trial.

112   Because we decoded and tracked the DV in real-time, we were able to terminate the visual stimulus

113   in a neurally contingent manner and probe both neural activity and the subject's behavior with

114   high precision and negligible latency (<34 ms, see Methods). Inspired by sequential sampling

115   behavioral models that assume a bound[12-14], the first closed-loop test we performed was to impose

116   virtual decision boundaries that, if reached, would result in stimulus termination (Fig. 2a),

117   prompting the subject to immediately report its decision (in trials with no delay period). In this

118   manner we obtained a direct mapping between the nearly instantaneous readout of decision state

119   and the likelihood of a given behavioral choice.

120   Figure 2b shows 22 example DV traces from trials that led to stimulus termination by reaching a

121   fixed DV boundary of magnitude 3, within a tolerance of ± 0.25 DV units.

122   To characterize the relationship between the DV at termination and prediction accuracy, we

123   systematically swept the parameter space for the boundary height using values spanning 0.5-5 DV

124   units in 0.5 increments (1DV unit corresponds to an increase of 2.718 in the likelihood ratio of

125   choosing one target over the other). Figure 2c shows that prediction accuracy increases

126   monotonically with the DV magnitude at termination as expected. Moreover, using only 100 ms

127   of data to estimate the DV that triggered termination, the difference between the observed

128   likelihood of a given choice (solid trace) and that predicted by the logistic function (dashed trace)

129   was only, on average, 1.7% (1.9%) for monkey H (F) (Fig. 2c, Extended Data Fig. 2c). For

130   example, neural DV values of ±3 predict decisions upon termination with an accuracy of 98%.

131   Even DV values as low as ±0.5-1 predict decisions with an accuracy of nearly 70%. DV

132 fluctuations below ±0.5 are more susceptible to noise in our estimates of decision state and at most

133 are associated with very weak choice preferences and were thus not tested. Overall, these results

134 show that moment-by-moment fluctuations in PMd/M1 neural population activity captured by our

135 decoding model are indeed reflective of a fluctuating internal decision state of the animal—

136 fluctuations that have been covert and thus uninterpretable until now.

137 Figure 2c (Extended Data Fig. 2c) combines trials across a wide range of coherences and stimulus

138 durations, aggregated across 17 (15) sessions from monkey H (F). To identify experimental factors

139 that might influence the observed relationship between DV at termination and prediction accuracy,

140 we first resorted the same trials in Figure 3c by stimulus coherence. The results show that there is

141 a small separation between the curves for high and low coherence trials (Fig. 2d) with higher

142 accuracy for high coherence trials. The shift is small but reliable across monkeys (Extended Data

143 Fig. 2d). We hypothesized that this difference resulted from motion energy signals already en route

144 from the retina to PMd/M1 (~175 ms latency) when the DV reached stimulus termination. More

145 motion energy signals would be arriving from this neural 'pipeline' on high coherence trials,

146 leading to a slightly higher DV than we measured at stimulus termination.

147 To assess this possibility, we measured the derivative of the DV around termination and performed

148 the following two analyses. First we checked whether DV derivative explained a significant

149 fraction of choice variance beyond DV value alone (see Methods). For both monkeys the effect of

150 DV derivative (defined as the DV slope in the last 50 ms of stimulus presentation) was significant

151 ($p=0.02$, $p = 4.5 \times 10^{-11}$ for monkey H and F, respectively) and the effect was congruent with our

152 hypothesis: stronger positive derivatives predicted higher likelihood of rightward choices and

153 stronger negative derivatives predicted higher likelihood of leftward choices (Extended Data Table

154 1, "DV diff"). Second, we tested whether high coherence trials were associated with higher DV

7

155   derivatives at termination by performing linear regression of DV derivatives as a function of signed

156   coherence. For both monkeys signed coherence was strongly predictive of DV slopes: $p = 2.17$

157   $\times 10^{-171}$ and $R^2 = 0.23$ for monkey H and $p = 1.57 \times 10^{-105}$ and $R^2 = 0.16$ for monkey F. These results

158   confirm that DV derivative is predictive of choice beyond DV alone and show that higher

159   coherence trials are associated with higher DV derivatives. The data are consistent with our

160   hypothesis above that the DV continues to evolve under the influence of 'pipeline' sensory

161   information for a short interval following stimulus termination, resulting in somewhat better

162   prediction accuracy than expected from the DV at termination, especially at high coherences.

163   Sorting trials by duration (Fig. 2e, Extended Data Fig. 2e), reveals a different effect: the centers of

164   the quantiles are strongly shifted to the right (higher DV magnitudes) for longer stimuli compared

165   to shorter stimuli. This effect is expected from multiple sequential sampling models[8,15-17]. In drift

166   diffusion models, for example, diffusion to high decision bounds requires more time than for low

167   bounds[18]. However, we tested whether stimulus duration *per se* was a significant predictor of

168   choice independently of DV value by including two additional regressors in our logistic model of

169   choice: stimulus duration (representing choice bias as a function of time) and an interaction term

170   between stimulus duration and direction (representing increased sensitivity to stimulus coherence

171   as function of time). Neither regressor was significant for either monkey ($p>0.05$, Extended Data

172   Table 1), implying that the likelihood of making one or the other choice depended on DV value

173   independently of the time required to reach that value.

174   Together, these results show that fluctuations in DV magnitude at a 100 ms time scale have a

175   predictable correlate in choice likelihood that is lawfully influenced by stimulus coherence and

176   robust across time. We emphasize that our decoded DV is model-based and thus a proxy for the

177   actual decision state in the brain. We are sampling from a relatively small number of neurons, and

8

178   the underlying mechanism is unlikely to be strictly linear (in contrast to the logistic model). In

179   addition, we do not know with certainty when the deliberation process ends within the brain, which

180   could occur before or after our stimulus termination on individual trials. Despite these caveats, our

181   ability to predict choice likelihood using a DV boundary criterion at stimulus termination within a

182   very small margin of error (<2% on average) confirms that DV is a reliable proxy for decision

183   state.

184   **Neurally detected CoMs can be validated and match the statistical regularities expected from**

185   **previous studies**

186   The mapping between DV and choice likelihood obtained in the first experiment (Fig. 2c), enabled

187   us to perform a new closed-loop experiment aimed at capturing particularly robust DV fluctuations

188   in which the sign of the DV (and thus the neurally inferred decision state of the animal) changed

189   in the middle of a trial, suggestive of a 'change of mind' at the behavioral level (CoM, Fig. 3a-b).

190   When the neural criteria for a CoM were met in real-time (see Methods, examples in Figure 4a,

191   orange and green arrows), the stimulus was terminated instructing the monkey to make a decision

192   as described above. Our aim was to detect neurally-based candidate CoMs, assess the influence

193   of the decision states before and after the CoM on the final choice, and determine whether

194   statistical properties of the neurally derived CoMs match the properties expected of CoMs from

195   prior psychophysical and neurophysiological studies.

196   We conceptually divide a CoM trial into two segments—the initial preference prior to the DV sign

197   change, and the final (opposite) preference that leads to the observed choice. The observed choices

198   allow corroboration of the neural estimate of the final decision state in the second segment

199   (Extended Data Fig. 6). For monkey F, the relationship between choice prediction accuracy and

9

200   DV at stimulus termination for CoM trials was very similar to that of non-CoM trials (compare

201   Extended Data Fig. 2c and Extended Data Fig. 6b, mean error between predicted and observed

202   choice likelihood: 1.9% for non-CoM trials vs 3.8% for CoM trials). This relationship was lawful

203   and monotonic for monkey H as well although lower than expected (Extended Data Fig. 6a,

204   compared to Fig. 2c, mean error between predicted and observed choice likelihood: 1.7% for non-

205   CoM trials vs 9.3% for CoM trials), suggesting that in addition to the measured DV at stimulus

206   termination, monkey H's decisions were also influenced by some aspect of the DV trajectory

207   history specifically related to the CoM. We formally tested this hypothesis by regressing choice as

208   a function of 3 additional parameters (in addition to DV at termination) that were enforced and

209   monitored in this experiment (see Methods): maximum DV deflection before sign change and

210   duration of sign stability before and after DV sign change. For monkey F, no additional factor was

211   choice predictive, whereas for monkey H both the duration of sign stability before and after the

212   CoM were also choice predictive (Extended Data Table 2) as suspected from Extended Data Fig.

213   6.

214   We combined all 985 (1727) CoM's detected in monkey H (F) to assess whether our neurally

215   detected CoMs conformed to three statistical regularities of CoMs established in previous

216   psychophysical[7] and electrophysiological[1] studies.

217   The first observation is that CoMs are more frequent for low and intermediate coherence trials as

218   opposed to high coherence trials, as high coherences are more likely to lead to straightforward

219   integration of evidence toward the correct choice. We found the same to be true in our real-time

220   detection data (Fig. 3c, Extended Data Fig. 2f; linear regression $p<0.001$).

221   The second observation is that CoMs are more likely to be corrective than erroneous. This

222   prediction results from the corrective role of additional visual evidence on the initial preference of

223   the subjects. This trend was also verified in the CoMs we detected with CoMs for non-zero

224   coherences and for both monkeys being more likely corrective than erroneous (Fig. 3d, Extended

225   Data Fig. 2g; Wilcoxon rank sum test p<0.001, median corrective and erroneous CoM counts: 530

226   and 242 for monkey H and 1046 and 443 for monkey F, respectively).

227

228   Finally, the third observation made in these previous studies was that CoMs were more frequent

229   early in the trial than later in the trial, consistent with drift diffusion models in which the DV is

230   more likely to have hit an absorbing decision bound as the trial progresses. We observed this effect

231   in our real-time, neurally detected CoMs as well (Fig. 3e, Extended Data Fig. 2h).

232   We also discovered a new regularity associated with CoMs: the average time of zero crossing was

233   negatively correlated with stimulus coherence (Fig. 3f, Extended Data Fig. 2i). This observation

234   likely results from the stronger corrective effect of higher coherence stimuli (Fig. 3d, Extended

235   Data Fig. 2g).

236   Together, these results show that robust fluctuations in DV that imply a change in choice

237   preference (zero crossing) can be captured in real time and validated as changes of mind.

238   **Pulses of additional visual motion evidence have smaller neural and behavioral effects when**

239   **presented at larger DV values**

240   In a final set of closed-loop experiments whether the neural and behavioral responses to brief

241   pulses of additional motion information varied with the state of the DV before the pulse. Inspired

242   by decision-making models involving buildup of neural activity to a bound[15,16,19,20], we expected

11

243   termination of the deliberation process and commitment to a choice to be more likely at high DV

244   values[1,7,8,16,21]. We therefore hypothesized that additional pulses of sensory evidence would result

245   in less change in DV and behavior when pulses were triggered by high DV values.

246

247   To characterize the relationship between DV and responses to a stimulus pulse, we again imposed

248   virtual DV boundaries (as in Fig. 3a-b) that, if reached, triggered a 200-ms pulse of additive dots

249   coherence (randomly assigned to be rightward or leftward on each trial) followed by stimulus

250   termination (Fig. 4a). We swept a subset of the previously used DV values for the boundary

251   (spanning 1-4 DV units, in 1.0 increments). Pulse strength was calibrated to yield very small but

252   significant effects on behavior, in an effort to avoid making the pulses so salient as to change the

253   animals' integration strategy on pulse trials (Δcoherence = 2% for monkey H, 4.5% for monkey

254   F). Pulse information had no bearing on the reward[8,17]. Motion pulses slightly but significantly

255   biased the monkeys' choices in the direction of the pulse (p = 8.38E-14 for monkey H, Fig. 4b; p

256   = 1.95E-4 for monkey F, Extended Data Fig. 2j).

257

258   We reasoned that, to detect the presumably small effects of these small motion pulses on the DV,

259   we would need to account for a processing delay for changing stimulus information to influence

260   our recorded neural populations. Thus, to quantify the effect of the pulse on the evolving DV, we

261   first measured the minimum latency for visual stimulus information to influence the DV: we

262   calculated the time after stimulus onset at which the DV traces diverged for rightward vs. leftward

263   choices in an independent set of open loop trials at the strongest motion coherence. We refer to

264   this time point as the evidence representation latency (ERL). For each trial, we measured the

265   change in DV (ΔDV) for each time bin, beginning at the time of pulse onset plus the ERL (or

266   PERL—see Methods). We found that, on average, motion pulses slightly but significantly biased

267   ΔDV in the direction of the pulse (Fig. 4c, Extended Data Fig. 2k).

268

269   In the case of simple, unbounded linear integration, we expect the magnitude of DV change in

270   response to a fixed motion pulse to remain constant regardless of the triggering DV at pulse onset.

271   In contrast, Fig. 4d (Extended Data Fig. 2l) shows that motion pulses led to larger DV changes

272   when triggered by low as compared to high DV values.

273

274   Previous studies have shown that behavioral and LIP neural responses to similar motion pulses

275   tend to be smaller when pulses are delivered later in the stimulus[8,17]. Large DV values tend to

276   occur later in the trial, and this was hypothesized to be the underlying reason for the diminishing

277   pulse effects (assuming some sort of bound on integration of evidence at larger DVs); but these

278   studies lacked concurrent neural population recording and decoding and thus did not have access

279   to the momentary decision state. Thus, the time of pulse onset is a possible confound for the

280   decreasing pulse effects at high DV bound values as depicted in Fig 4d. To control for this

281   possibility, we first used the slope of the ΔDV vs. time relationship measured on individual trials

282   (ΔDV slope) to summarize the effect of the stimulus pulse on DV on single trials. We then

283   performed a multiple regression analysis of ΔDV slope that included both the triggering DV value

284   and the time of the motion pulse as regressors (and other variables as well - see Methods). The

285   regression data confirm that the effect of stimulus pulses on DV is only significant when triggered

286   by lower DV values, and that this effect is not explained by pulse timing (Fig. 4e, Extended Data

287   Fig. 2m, Supplementary Information Table 1). Similarly, the effect of motion pulses on

288   psychophysical behavior is weaker when triggered by high DV values, and these effects also are

13

289   not explained by pulse timing (Fig. 4f, Extended Data Fig. 2n, Supplementary Information Table

290   2).

291

292   Our finding that larger DVs (and corresponding behavioral readouts) are more resistant to pulses

293   is consistent with several models of decision formation, including linear integration to a decision

294   bound (such as a simple stopping criterion[12]) or a more complex nonlinear integration

295   process[17,22,23]. Inspired by these results, we returned to the data from the first two experiments in

296   an attempt to explore the nature of the apparent bounding mechanism by analyzing the time-

297   variance of the DV over the course of individual trials. In the case of an absorbing decision bound

298   or attractor network, we would expect DV variability (measured as the DV time derivative) to

299   decrease after reaching the bound. We indeed found that, on average, DV variability decreases

300   over time within single trials (Fig. 5a, Extended Data Fig. 7a). This effect holds across all stimulus

301   strengths, although variability peaks earlier and falls faster on the easiest trials (Fig. 5b, Extended

302   Data Fig. 7b).

303

Discussion

305   While previous single-electrode recordings have strongly advanced our understanding of the

306   neural correlates of perceptual decision-making, interesting dynamics in choice signals were lost

307   to necessity of averaging data across trials. With a few notable exceptions[3,24], deploying the

308   statistical power of simultaneous multi-electrode recordings to track single-trial population

309   dynamics during choice behavior is a recent advance[1,2,6]. Even more recent work has leveraged the

310   power of brain-computer interfaces (BCIs) to study neural correlates of prediction, learning, and

311   multisensory integration (as reviewed in Golub et al. 2016[25]). In this study, for the first time, we

14

312   probed moment-to-moment fluctuations in decision states using BCI-inspired closed loop

313   experiments that enabled neurally contingent stimulus control and made behavioral validation of

314   these fluctuations feasible (see Methods). We show that large fluctuations (up to several log units)

315   in a decoded decision variable in premotor and primary motor cortices are nearly instantaneously

316   (<100 ms) predictive of choice. We captured neural correlates of changes of mind in the form of

317   robust changes in DV sign. The statistical regularities of these rare events match previous

318   psychophysical CoM findings. Finally, we showed that larger DV values are resistant to additional

319   pulses of sensory evidence, supporting the hypothesis that large DVs are associated with higher

320   commitment to an upcoming choice.

321

322   Importantly, the impressive choice prediction accuracy achieved in this study using a linear

323   decoder does not imply that the brain's decision formation process is also linear. In principle, such

324   a decoder could predict binary choices quite well even if the true neural process underlying

325   decision formation were nonlinear, depending on the form of the nonlinearity (see, e.g., Sussillo

326   et al. 2016[26] for an example of a linear neural to kinematic decoder which only slightly

327   underperforms a more powerful nonlinear recurrent neural network). However, our linear DV is

328   tightly linked to choice behavior (e.g. Fig. 2c), showing that variations in DV magnitude

329   meaningfully track the ongoing process of decision formation despite the possible presence of

330   nonlinearities in the underlying neural mechanism.

331

332   Previous studies have described fluctuations in offline decoded decisions associated with changes

333   of mind[1-3,6]. Here we confirm and extend those observations with neurally contingent interrogation

334   of candidate CoM events, but we also find large, behaviorally relevant fluctuations even when the

335   DV remains on one side of the discriminant hyperplane in non-CoM trials (e.g. Figs. 1e and 2b).

336   We wondered whether these DV fluctuations were related to stochastic variations in motion

337   strength of the stimulus on single trials. While across coherence levels the average motion energy

338   explains a large portion of DV variance (Extended Data Fig. 8a-b), our data shows that within

339   coherence stochastic fluctuations in the stimulus are not the dominant cause of DV fluctuations

340   (Extended Data Fig. 8c-d). Further experiments will be needed to address the source(s) of these

341   fluctuations and their relationship with fluctuations in other brain areas[27] as well as other cognitive

342   processes including motor preparation and execution[28,29], attention, motivation, and confidence.

343

344   In addition to validating the behavioral relevance of neurally detected DV fluctuations, our ability

345   to impose real-time task changes contingent upon them allowed us to show that neural and

346   behavioral responses to pulses of additional sensory evidence diminish when pulses are presented

347   at larger momentary DV values. These results, combined with the reduction in DV variability

348   observed over the course of single trials, suggest the presence of an absorbing decision bound in

349   these motor cortical neural populations, consistent with attractor dynamics in which the neural

350   population converges on a stable state as a decision is formed[22,23].

351

352   The conceptual and technical innovation that enabled these findings is our ability to accurately

353   decode decision states in real time, which could bring the concept of cognitive prostheses[30-33] much

354   closer to reality by providing another means of decoding subjects' goals for use as a flexible

355   prosthetic control signal. More broadly, the real-time closed loop approach demonstrated here may

356   be applicable not only to decision-making processes, but also to other cognitive phenomena such

357   as working memory and attention.

16

358
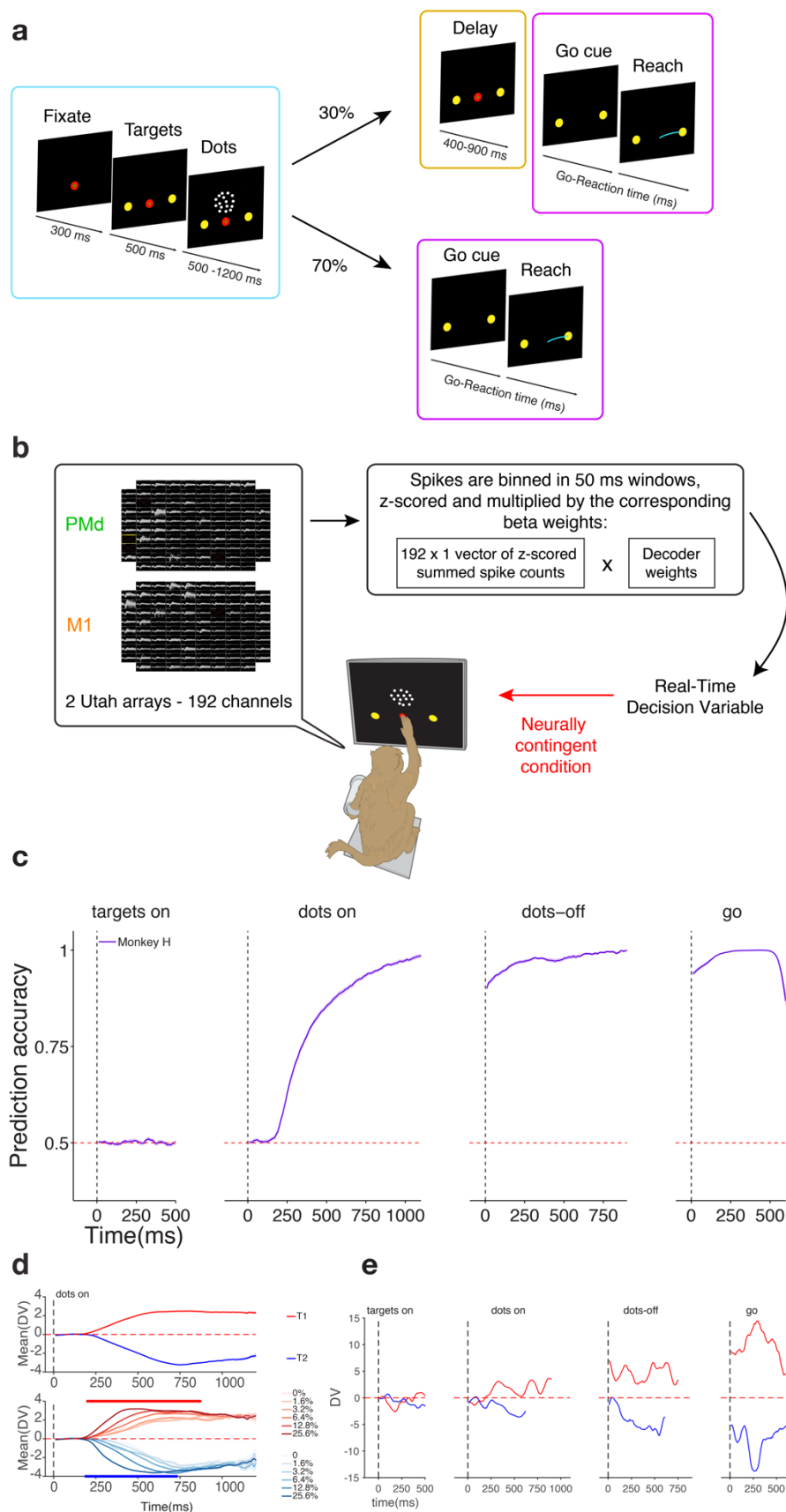
359   **Figure 1. Setup and performance of real-time readout of decision states during a motion**

360   **discrimination task.**

361

362   **a) Motion discrimination task.** Trials began with the onset of a fixation point (FP) on the

363   touchscreen. Once both eye and hand fixation were acquired, two targets appeared on the screen.

364   The motion stimulus was shown after a short delay (500 ms) and lasted 500-1200 ms for the open-

365   loop trials. On 70% of the trials the dots offset was followed by the go cue (no delay period), while

366   on the remaining 30% the subject was required to withhold a response for a random delay duration

367   (400-900 ms). Decision states were continuously decoded during all epochs of the trial. Three

368   different decoders were used during different trial epochs, shown by the different colored boxes

369   (blue, yellow and purple; see Methods).

370   **b) Real-time, closed-loop setup.** Neural activity from 96-channel Utah Arrays was continuously

371   recorded and processed while monkeys performed the motion discrimination task. For monkey H,

372   two Utah arrays implanted in PMd and M1 were used. For monkey F only one Utah array

373   implanted in PMd was utilized. During data collection, the recorded neural activity was binned,

374   summed, z-scored and projected onto a single dimension: a linear choice decoder. The result of

375   this operation was our real time read out of commitment, which could be used to stop the stimulus

376   presentation in a neurally contingent manner (red arrow), thereby closing the loop in the

377   experiment.

378   **c) Choice prediction accuracy obtained from real-time, open-loop readout.** Average

379   prediction accuracy (see Methods) over time ± SEM for monkey H is plotted in purple. Prediction

380   accuracy is calculated for each time point aligned to four different events in the trial (targets onset,

381   dots onset, dots offset and go cue) using the real-time DV and quantified as the fraction of trials in

18

382 which the classifier correctly predicted the monkey's upcoming choice. For logistic regression this

383 operation is equivalent to comparing the DV sign to the choice sign. Accuracy was calculated for

384 each session and averaged across sessions using a total of 16468 trials for monkey H.

385 **d) Average Decision Variable traces during dots period.** Top panel: Average DV during the

386 dots epoch for right (red) and left (blue) choices for monkey H. Bottom panel: Average DV sorted

387 by choice and stimulus coherence (correct trials only) for monkey H. Darker shades correspond to

388 higher stimulus coherence. Red and blue dots indicate timepoints for which coherence was

389 significant regressor of DV for T1 and T2 choices respectively (correct trials only, $p<10^{-5}$

390 uncorrected). For monkey H coherence is a significant regressor of DV for at least one of the

391 choices for the period between [190, 870] ms aligned to dots onset.

392 **e) Example DV traces captured during open loop trials.** DV traces for two trials are plotted as

393 a function of time aligned to four different events: targets onset, dots onset, dots offset and go cue.

394 The trial in red led to a right choice whereas the trial in blue led to a left choice. Despite the stability

395 in DV sign for these two trials from ~250 ms after dots onset until the end of the trial, strong

396 fluctuations in DV magnitude were observed in both cases, within and across epochs.
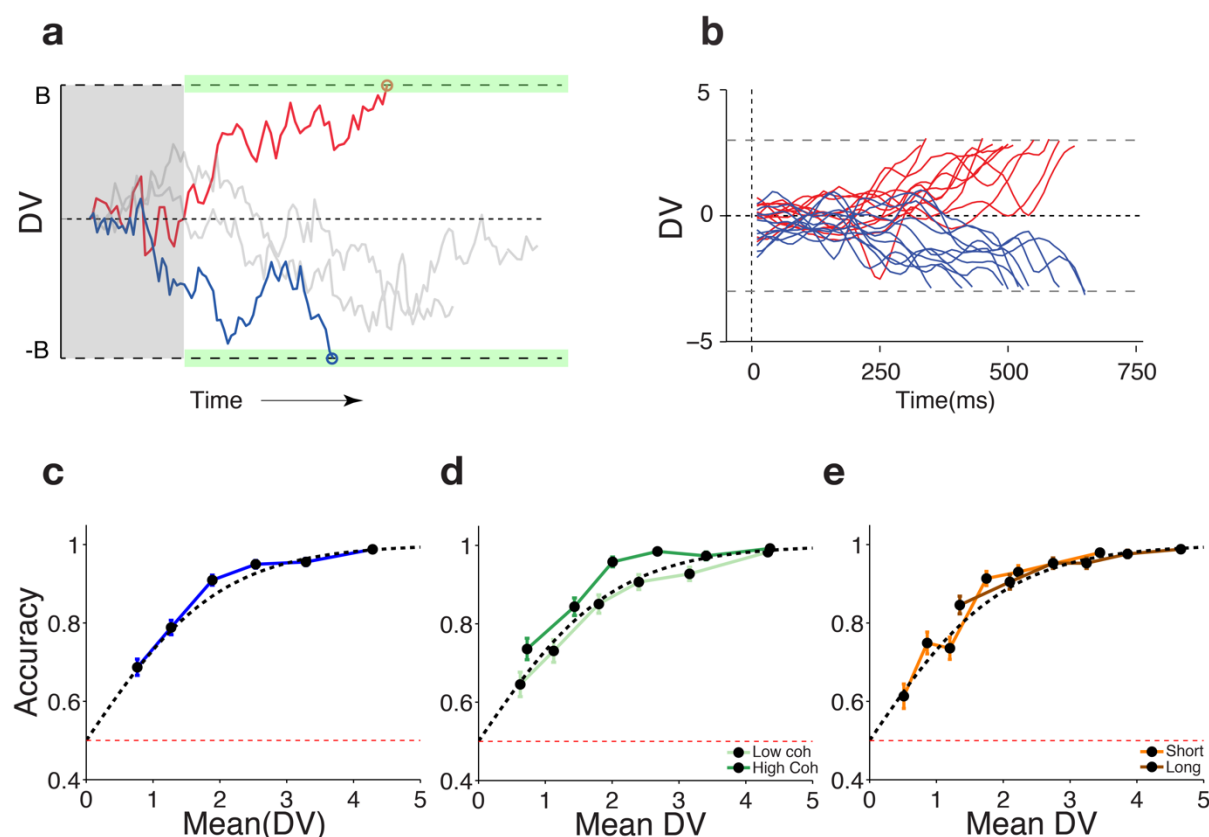
Figure 2. Choice likelihood can be accurately decoded in real-time across experimental conditions using only 50 ms of neural data.

a) Schematic of the first closed loop experiment implemented in real time. Virtual boundaries for DV magnitude (green shaded regions) were imposed and if reached, triggered the termination of the stimulus presentation. The subject was then immediately asked to report its decision. A 250 ms minimum stimulus duration was imposed (grey shaded region) to prevent random fluctuations in the beginning of the trial from triggering stimulus termination. If the boundary wasn't reached, the stimulus was presented for a pre-selected random duration (500-1200 ms). Grey traces show cartoons of trials for which the boundary was not reached while red (blue) traces show terminated

409   trials that the decoder predicted would result in a right (left) choice. 5 different boundary values

410   were used on each experiment.

411   **b) Example trials captured during the virtual boundary experiment.** Real-time DV time

412   courses for example trials terminated using boundaries set at $+3$ and -3 DV units. Traces are

413   colored according to behavioral choice at the end of the trial: right choices in red and left choices

414   in blue. Data from one session from monkey H.

415   **c) Prediction accuracy as a function of DV magnitude.** Choice prediction accuracy for all trials

416   collected during virtual boundary experiment as a function of DV magnitude for monkey H is

417   shown in blue. Trials were split in 6 quantiles sorted by DV magnitude at termination. Prediction

418   accuracy and median DV magnitude were calculated and plotted separately for each quantile (blue

419   line with black symbols). Blue error bars show standard error of the mean for a binomial

420   distribution. Dashed black line shows predicted accuracy from log-odds equation and red dashed

421   line shows chance level. Data from 2973 trials from monkey H.

422   **d) Prediction accuracy as a function of DV and stimulus coherence.** Same data shown in **c)** but

423   having pre-sorted the trials by coherence (see Methods). Dark green trace shows high coherence

424   results and light green, low coherence results. Same conventions as in **c).**

425   **e) Prediction accuracy as a function of DV and stimulus duration.** Same data shown in **c)** but

426   having pre-sorted the trials by stimulus duration. Brown trace shows results for long trials and

427   orange trace results for short trials (see Methods). Same conventions as in **c).**

428

429

430
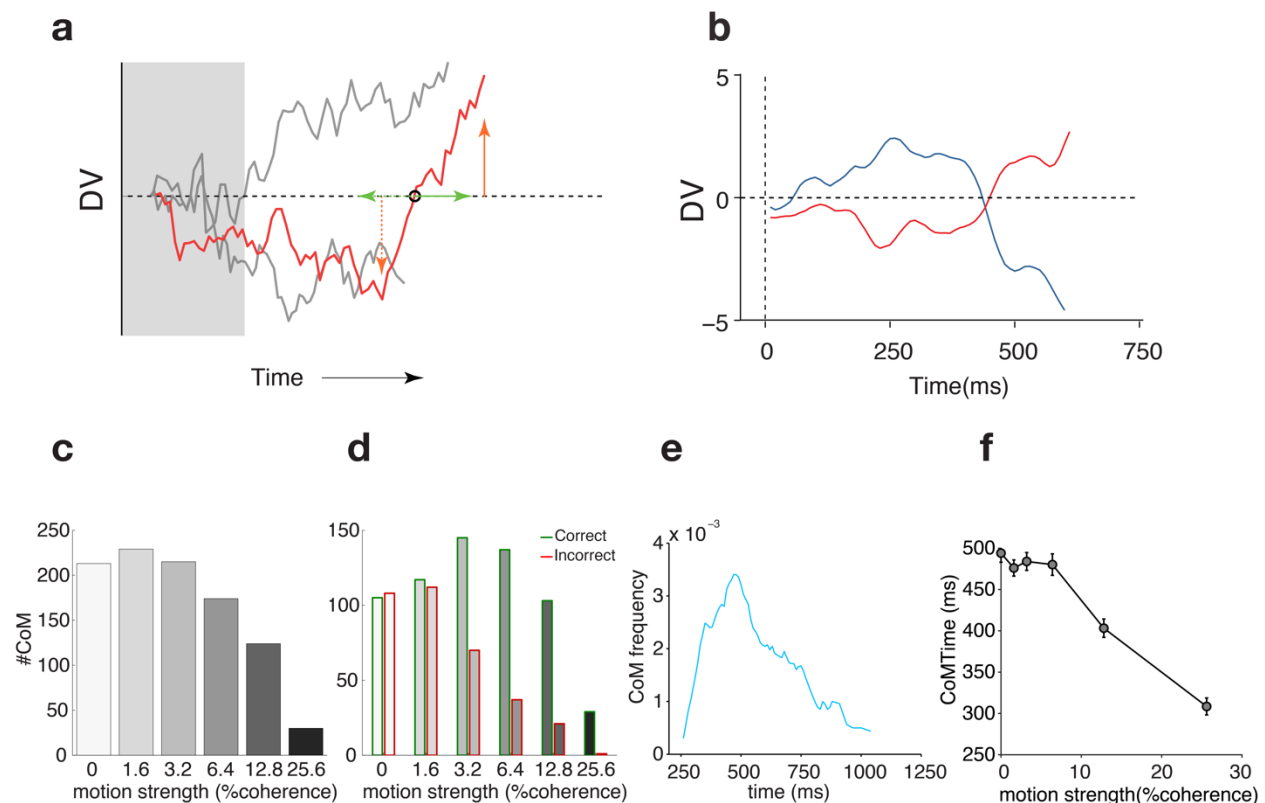
431   **Figure 3 Putative changes of mind can be detected and validated in real time.**

432

433   **a) Schematic of the second closed loop experiment implemented in real time.** The value and

434   history of the DV trace were tracked on each trial. If a 0 crossing (sign change in the DV) was

435   detected, the conditions required for termination were checked and termination was carried out if

436   the conditions were met (see Methods). In this example the conditions for temporal stability of DV

437   sign are depicted by the green horizontal arrows while the conditions for minimum DV deflection

438   before and after CoM are depicted by the orange arrows. Upon termination, the subject was

439   immediately asked to report its decision. A 250 ms minimum stimulus duration was imposed (grey

440   shaded region) such that random fluctuations in the beginning of the trial did not trigger stimulus

22

441 termination. If the conditions were not met or if a 0 crossing was never detected, the stimulus

442 would be presented for a pre-selected random duration (500-1200 ms). Grey traces show cartoons

443 of trials for which the 0 crossings would not meet the criteria while the red the trace shows a

444 terminated trial that was predicted to lead to a rightward choice. One set of criteria for CoM validity

445 was used in each session (Extended Data. Table 4).

446 **b) Example trials captured during the CoM experiment.** Real-time DV time courses for 2

447 example trials with a putative CoM terminated after conditions were met (minimum DV pre and

448 post CoM: 2 and minimum period of sign stability pre and post CoM: 150 ms). Traces are colored

449 according to behavioral choice at the end of the trial: right choices in red and left choices in blue.

450 Two trials from one session from monkey H.

451 **c) CoM frequency as a function of coherence**. Total number of CoMs detected for each

452 coherence for monkey H.

453 **d) CoM frequency as a function of coherence and direction**. Total number of CoMs detected

454 for each coherence and direction for monkey H. Red bars correspond to erroneous CoMs and green

455 bars to corrective CoMs.

456 **e) CoM frequency as a function of time in the trial**. Frequency of CoMs detected as a function

457 of time during stimulus presentation for monkey H. Because only CoMs that would have resolved

458 by 250 msec after stimulus onset were considered, there is an edge effect with CoM frequency

459 briefly increasing between ~250-450 msec after which it declines.

460 **f) CoM time as a function of coherence.** Average CoM time (defined as the zero crossing for

461 each CoM trial) is plotted as a function of stimulus coherence. Error bars show s.e.m across trials

462 for each condition. CoM time was negatively correlated with stimulus coherence (p =1.8 x 10^{-17})

463

23

465
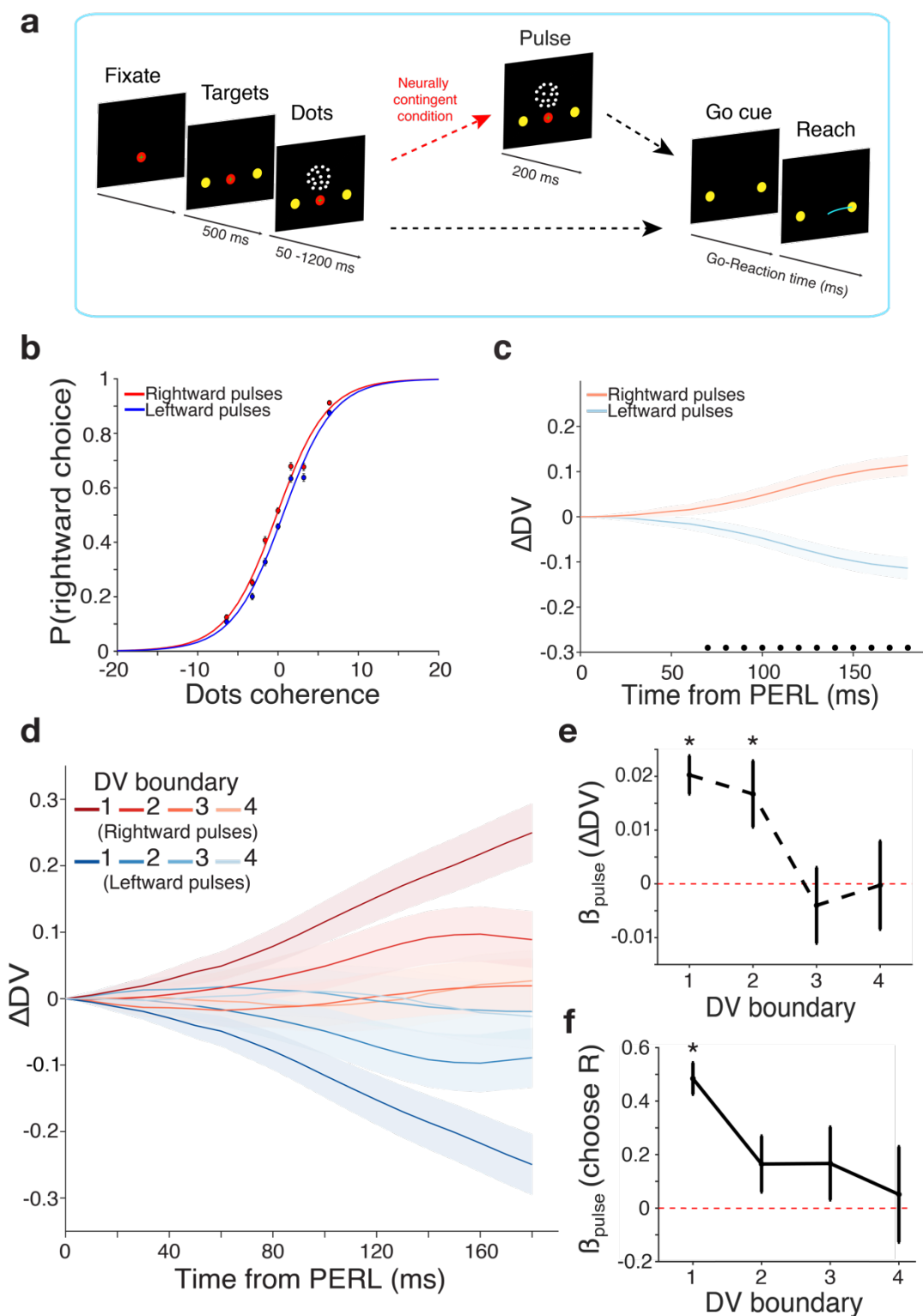
466   **Figure 4. Neurally triggered pulses of motion evidence nonlinearly bias both choice behavior**

467   **and DV.**

468

469   **a) Motion pulse task.** As in the motion discrimination task, trials began with the onset of a fixation

470   point (FP) on the touchscreen. Once both eye and hand fixation were acquired, two targets

471   appeared. The motion stimulus was shown after a short delay (500 ms) and a maximum stimulus

472   duration was randomly assigned from 500-1200 ms. Virtual boundaries for DV magnitude were

473   imposed (randomly assigned to integer values from 1-4) and if reached, triggered a 200-ms pulse

474   of additive dots coherence, randomly assigned to be rightward or leftward on each trial (± 2%

475   coherence for monkey H), followed immediately by termination of the dots stimulus. A 50 ms

476   minimum stimulus duration was imposed to ensure a minimum total stimulus duration of 250 ms.

477   If the DV boundary wasn't reached, the dots stimulus was presented for a pre-selected random

478   duration (500-1200 ms). Dots offset was followed by the go cue. Decision states were continuously

479   decoded using the dots period decoder during all epochs of the trial (blue box, see Methods).

480   **b) Psychometric functions for pulse trials.** Curves were fit using logistic regression on choice

481   with signed stimulus coherence and pulse direction as predictors, plus an intercept term. Data

482   points show mean proportion of rightward choices for each stimulus coherence, ± s.e.m. The pulse

483   effect is equivalent to changing the overall stimulus coherence by 0.384% (standard error

484   0.0514%, p = 8.38E-14). Data from 9614 rightward and 9523 leftward pulse trials from monkey

485   H.

486   **c) Average change in post-pulse DV from estimated Pulse Evidence Representation Latency**

487   **(PERL), mean subtracted.** ΔDV is the difference in the DV at each time point from the DV at

25

488   the PERL (170 ms for monkey H; see Methods). The mean ΔDV across pulse directions in each

489   time bin has been subtracted for visualization. Shaded error bars correspond to mean ± s.e.m. Black

490   dots indicate time bins in which ΔDV is significantly different for trials with pulses in opposite

491   directions (false discovery rate 0.05). Data from same trials as **b)**.

492   **d) Average change in post-pulse DV for each DV boundary, mean subtracted.** Conventions

493   as in **b)** but separated by the DV boundary triggering motion pulses in each direction. Darker colors

494   correspond to smaller DV boundary magnitudes. Data from monkey H, minimum 1507 trials per

495   condition shown.

496   **e) Pulse coefficients from linear regression on ΔDV slope for each DV boundary.** ΔDV slope

497   is the single-trial slope of the ΔDV from PERL to either the animal's median go-reaction time or

498   150 ms prior to movement onset, whichever came first (as shown in **b)**). Multiple linear regression

499   was performed separately on ΔDV slope for trials at each DV boundary with the following

500   predictors: signed dots coherence, pulse onset time, pulse direction, pulse onset time * pulse

501   direction, plus an intercept term. Data points and error bars represent the coefficient for pulse

502   direction for trials at each DV boundary, ± s.e.m.; asterisks denote significantly nonzero

503   coefficients at 95% confidence. Data from same trials as **d)**.

504   **f) Pulse coefficients from logistic regression on choice for each DV boundary.** Logistic

505   regression was performed separately on the probability of a rightward choice for trials at each DV

506   boundary with the following predictors: signed dots coherence, pulse onset time, pulse direction,

507   pulse onset time * pulse direction, plus an intercept term. Data points and error bars represent the

508   coefficient for pulse direction for trials at each DV boundary, ± s.e.m.; asterisks denote

509   significantly nonzero coefficients at 95% confidence. Data from same trials as **d)**.

510
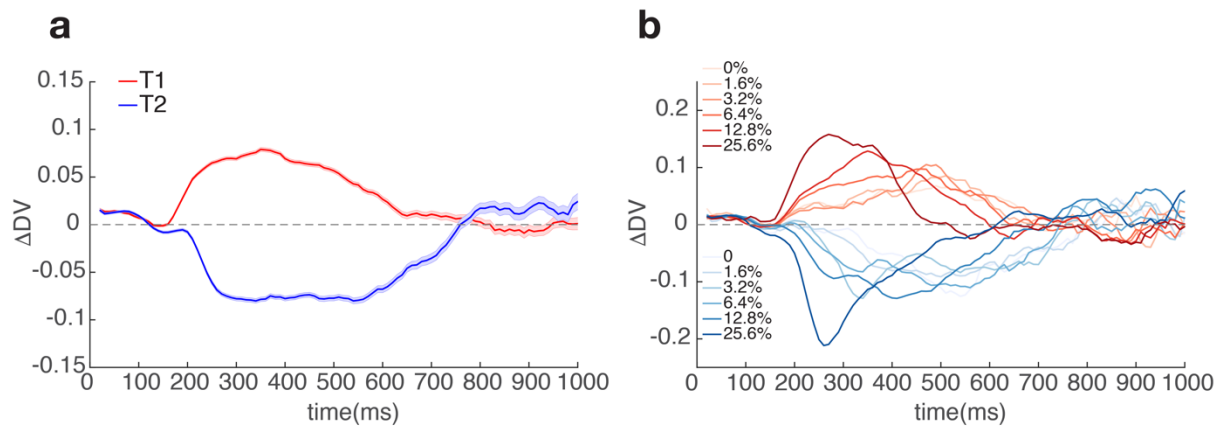
Figure 5. Within trial DV variability decreases over time for long duration stimuli.

a) Average DV derivative as a function of time and choice - monkey H. ΔDV was calculated for each trial as the difference between consecutive DV estimates spaced out by 10 ms. Traces show average ΔDV +/- s.e.m for right choices (red trace) and left choices (blue trace) during stimulus presentation. ΔDV initially starts increasing around the expected stimulus latency (170 ms) but progressively decreases for long (>600 ms) stimulus presentations.

b) Average DV derivative as a function of time and signed coherence - monkey H. Same data as in **a)** but with DV derivative averaged separately for each choice and motion coherence level (correct trials only). Right choices are plotted in red and left choices in blue as in **a)**. Darker traces correspond to stronger coherences.

27

Extended Data

Logistic Regression on Choice						
	Monkey H			Monkey F		
Predictor	Beta Value	95% CI	p-value	Beta Value	95% CI	p-value
Bias	-0.1613	[-0.2987 , -0.02381]	0.02147	-0.1197	[-0.2413 , 0.001806]	0.0535
Coherence	2.747	[2.268 , 3.227]	2.757e-29	1.885	[1.55 , 2.22]	2.912e-28
DV Termination	2.073	[1.779 , 2.367]	1.887e-43	1.708	[1.52 , 1.895]	1.626e-71
DV Diff	0.2989	[0.0488 , 0.5491]	0.01916	0.577	[0.4062 , 0.7478]	3.586e-11
Stimulus duration	0.1199	[-0.01489 , 0.2548]	0.08125	0.07474	[-0.04713 , 0.1966]	0.2293
Stimulus duration* Stimulus direction	-0.02803	[-0.2053 , 0.1492]	0.7566	0.0466	[-0.1132 , 0.2064]	0.5677

Extended Data Table 1 – Coefficients obtained from logistic regression on choice – virtual boundary experiment (monkeys H and F)

| Logistic Regression on Choice after CoM						
	Monkey H			Monkey F		
Predictor	Beta Value	95% CI	p-value	Beta Value	95% CI	p-value
Bias	-0.2876	[-0.4658 , -0.1093]	0.001568	-0.192	[-0.3773 , -0.00662]	0.04235
Coherence	1.304	[1.005 , 1.602]	1.189e-17	1.284	[0.96 , 1.608]	8.028e-15
DV Termination	1.67	[1.158 , 2.182]	1.623e-10	2.46	[1.872 , 3.049]	2.51e-16
DV max opposite	0.06967	[-0.4461 , 0.5854]	0.7912	-0.2387	[-0.8959 , 0.4185]	0.4766
Time after CoM * sign(DV Termination)	0.7065	[0.2317 , 1.181]	0.003544	0.02265	[-0.3835 , 0.4288]	0.913
Time before CoM * sign(DV max opposite)	0.7471	[0.3083 , 1.186]	0.0008467	-0.1427	[-0.6394 , 0.3539]	0.5732

28

534   **Extended Data Table 2– Coefficients obtained from logistic regression on choice - change of**

535   **mind experiment (monkeys H and F)**

536
537
538
539
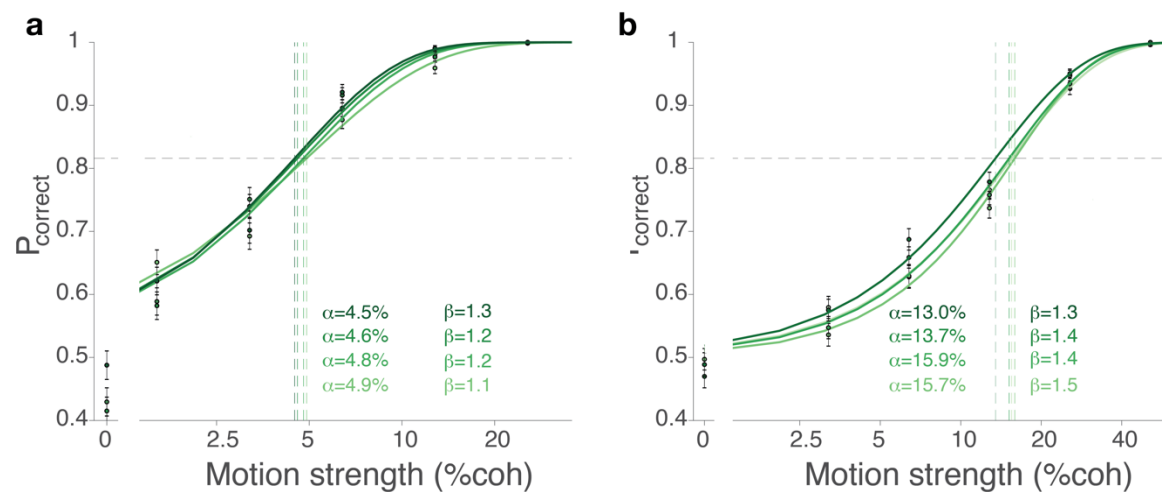
540
541
542   **Extended Data Figure 1 - Behavioral performance - Variable duration task.**

543

544   **a) Psychophysical performance for monkey H in the variable duration task.** Percentage

545   correct is plotted as a function of net motion coherence (calculated for both directions). Trials were

546   sorted for stimulus duration in 4 quartiles from long (dark green curve) to short (light green curve).

547   Data from each quartile were fit separately by a Weibull curve. Inset shows fit parameters for each

548   quartile. Data from 12516 open loop trials. Stimulus duration quartiles: Q1: [0.500 , 0.574] s Q2:

549   [0.574 , 0.680] s Q3[0.680 , 0.827] s Q4: [0.827, 1.200] s.

550   **b) Psychophysical performance for monkey F in the variable duration task.** Same as **a)** for

551   monkey F. Data from 12365 open loop trials. Stimulus duration quartiles: Q1: [0.500 , 0.574] s

552   Q2: [0.574 , 0.667] s Q3[0.667 , 0.813] s Q4: [0. 813, 1.200] s.

553

555
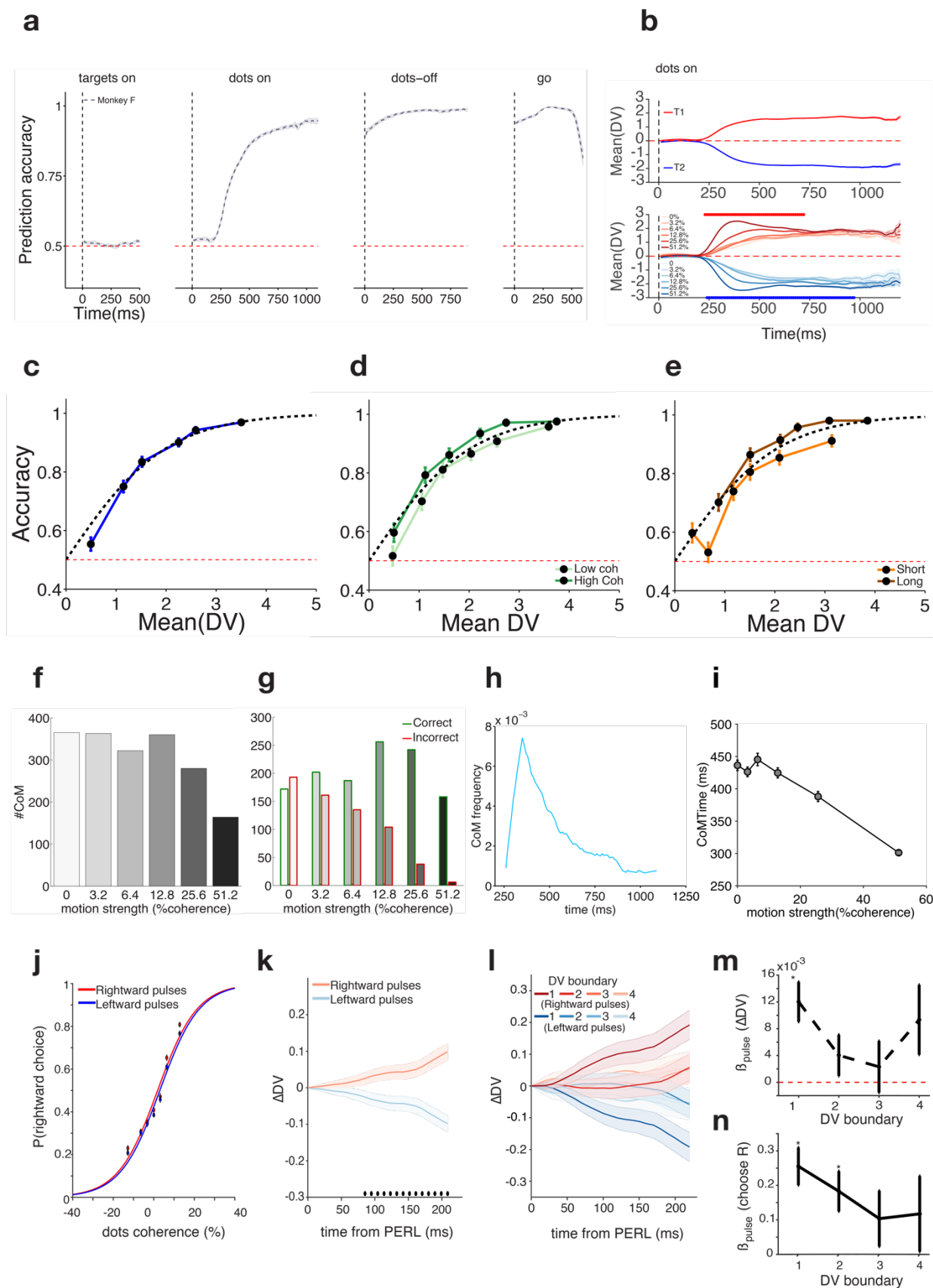
556 **Extended Data Figure 2 - Results for decoding and perturbation of DV in real time - monkey**

557 **F**

558

559 **a) Choice prediction accuracy obtained from real-time readout.** Same as **Figure 1c** for monkey

560 F. Accuracy was calculated for each session and averaged across sessions using a total of 15826

561 trials.

562 **b) Average Decision Variable traces during dots period.** Same as **Figure 1d** for monkey F. For

563 monkey F coherence is a significant regressor of DV for at least one of the choices for the period

564 between [230, 970] ms aligned to dots onset.

565 **c) Prediction accuracy as a function of DV magnitude.** Same as **Figure 2c** for monkey F. Data

566 from 2518 trials from monkey F.

567 **d) Prediction accuracy as a function of DV and stimulus coherence.** Same data shown in **c** but

568 having pre-sorted the trials by coherence (see Methods). Dark green trace shows high coherence

569 results and light green, low coherence results. Same conventions as in **c.**

570 **e) Prediction accuracy as a function of DV and stimulus duration.** Same data shown in **c)** but

571 having pre-sorted the trials by stimulus duration. Brown trace shows results for long trials and

572 orange trace results for short trials (see Methods). Same conventions as in **c).**

573 **f) CoM frequency as a function of coherence**. Same as **Figure 3c** for monkey F.

574 **g) CoM frequency as a function of coherence and direction**. Same as **Figure 3d** for monkey F.

575 **h) CoM frequency as a function of time in the trial**. Same as **Figure 3e** for monkey F.

576 **i) CoM time as a function of coherence.** Same as **Figure 3f** for monkey F. CoM time was

577 negatively correlated with stimulus coherence (p $=3.0$x 10^{-30})

578   **j) Psychometric functions for pulse trials.** Same as **Figure 4b** for monkey F. The pulse effect is

579   equivalent to changing the overall stimulus coherence by 0.545% (standard error 0.146%, p =

580   1.95E-4). Data from 10370 rightward and 9967 leftward pulse trials.

581   **k) Average change in post-pulse DV from estimated Pulse Evidence Representation Latency**

582   **(PERL), mean subtracted.** Same as **Figure 4c** for monkey F. PERL = 180 ms. Data from same

583   trials as **j).**

584   **l) Average change in post-pulse DV for each DV boundary, mean subtracted.** Same as **Figure**

585   **4d** for monkey F. Minimum 1731 trials per condition shown.

586   **m) Pulse coefficients from linear regression on ΔDV slope for each DV boundary.** Same as

587   **Figure 4e** for monkey F.

588   **n) Pulse coefficients from logistic regression on choice for each DV boundary.** Same as **Figure**

589   **4f** for monkey F.

590

591

33
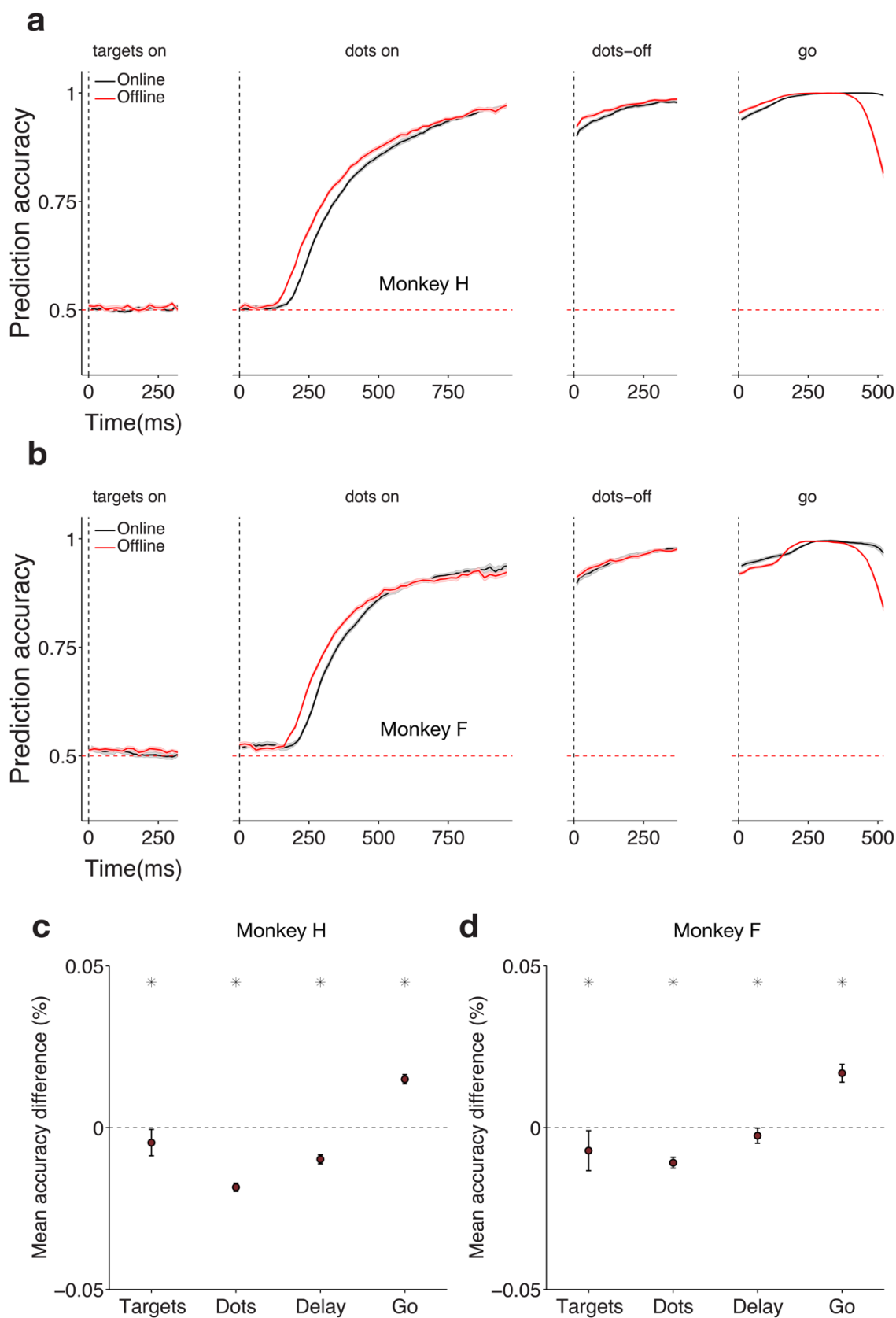
592 **Extended Data Figure 3 - Prediction accuracy online, offline and as a function of stimulus**
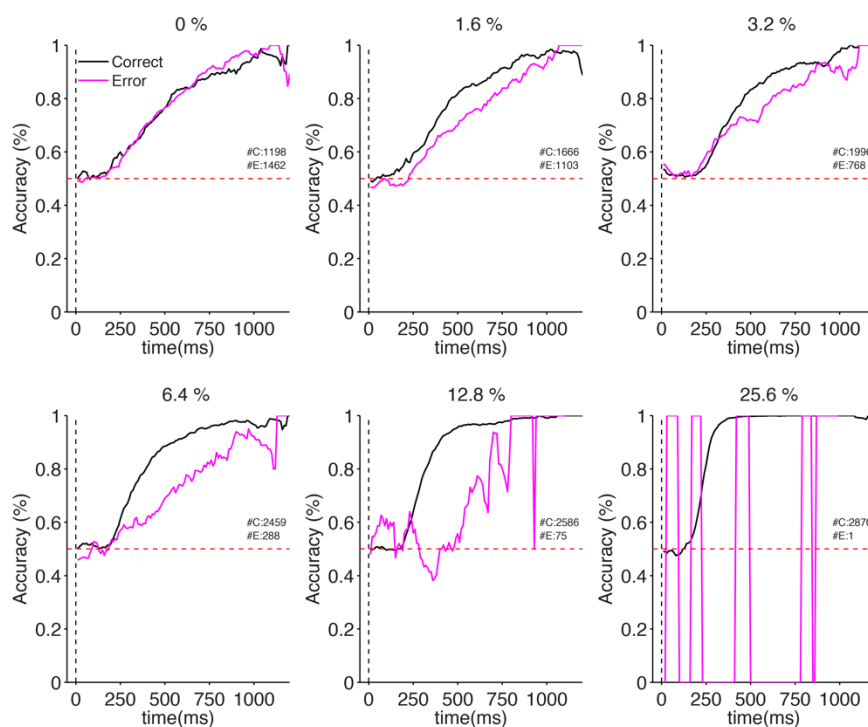
593 **coherence.**

594

595 **a) Online and Offline classifiers result in similar performance for targets, dots delay and**

596 **post-go epochs for monkey H.** Average prediction accuracy (see Methods) over time ± SEM

597 (across sessions) for monkey H. Online / offline classifier results are plotted in black / red. Data

598 in black is re-plotted from Figure 2a. Prediction accuracy is very similar online and offline across

599 the trial (see **c)**).

600 **b) Online and Offline classifiers result in similar performance for targets, dots delay and**

601 **post-go epochs for monkey F.** Same as **a** but for monkey F. Same conventions apply.

602 **c) Summary of performance difference between online and offline classifiers within each**

603 **epoch for monkey H.** Average performance difference between online and offline classifiers

604 (accuracy difference in percentage correct) for each of the epochs plotted in a). Positive number

605 numbers correspond to better online classifier performance and negative numbers to better offline

606 classifier performance. Black asterisks correspond to windows for which the differences were

607 significantly larger than zero (Wilcoxon signed-rank test, P<0.001).

608 **d) Summary of performance difference between online and offline classifiers within each**

609 **epoch for monkey F.** Same as **c)** for monkey F. For both monkeys **c)** and **d)** the difference of

610 choice prediction accuracies between the online and the offline classifiers was small and negative

611 for the target, dots and delay epochs (between -0.2% and -1.9%). In contrast, for the post-go

612 period, the difference in prediction accuracies was slightly positive (1.5% and 1.7% for monkey

613 H and F respectively).

614

Monkey H

Monkey F

615
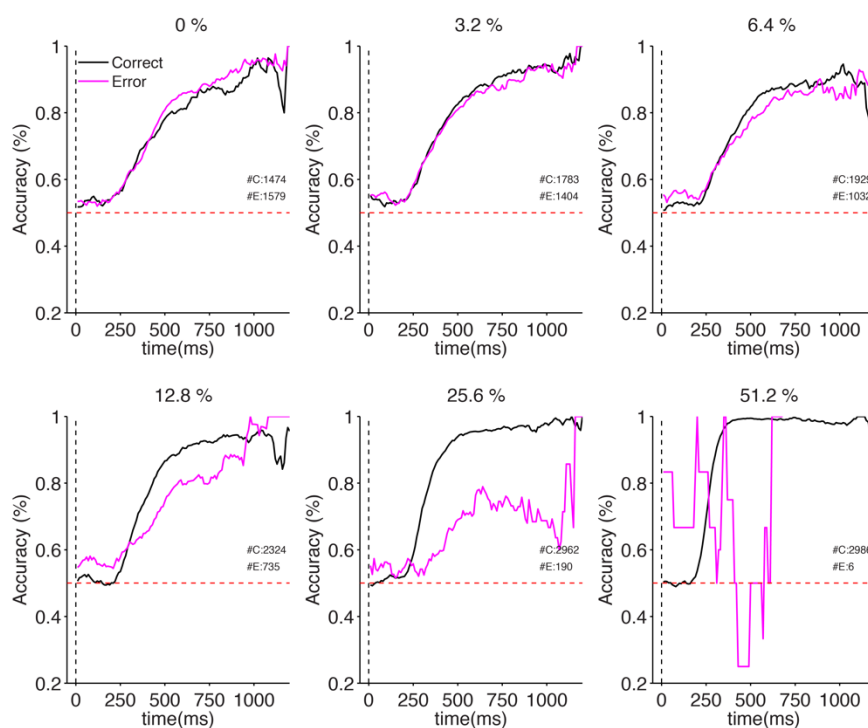
616 **Extended Data Figure 4 – Choice prediction accuracy for correct and incorrect trials as a**

617 **function of coherence.**

618

619 Choice prediction accuracy obtained from real-time readout for correct and incorrect trials for each

620 level of coherence. Prediction accuracy during dots epoch for each coherence level is plotted for

621 correct (black) and error (magenta) trials. Red dashed line corresponds to chance level. Insets show

622 total number of Correct (C) and Error (E) trials used in the analysis. Data for monkey H and F are

623 shown in top and bottom panels, respectively. Mean prediction accuracy for error trials after neural

624 latency (180 ms after stimulus presentation) is outside (and lower than) the 95% CI for correct

625 trials for 1.6%, 3.2%, 6.4%, 12.8% and 25.6% coherences for monkey H and for 12.8%, 25.6%

626 and 51.2% coherences for monkey F - 1000 bootstrap iterations. Results for the highest coherence

627 for each monkey should be interpreted carefully due to the extremely low number of error trials

628 for these conditions resulting from excellent behavioral performance.

629
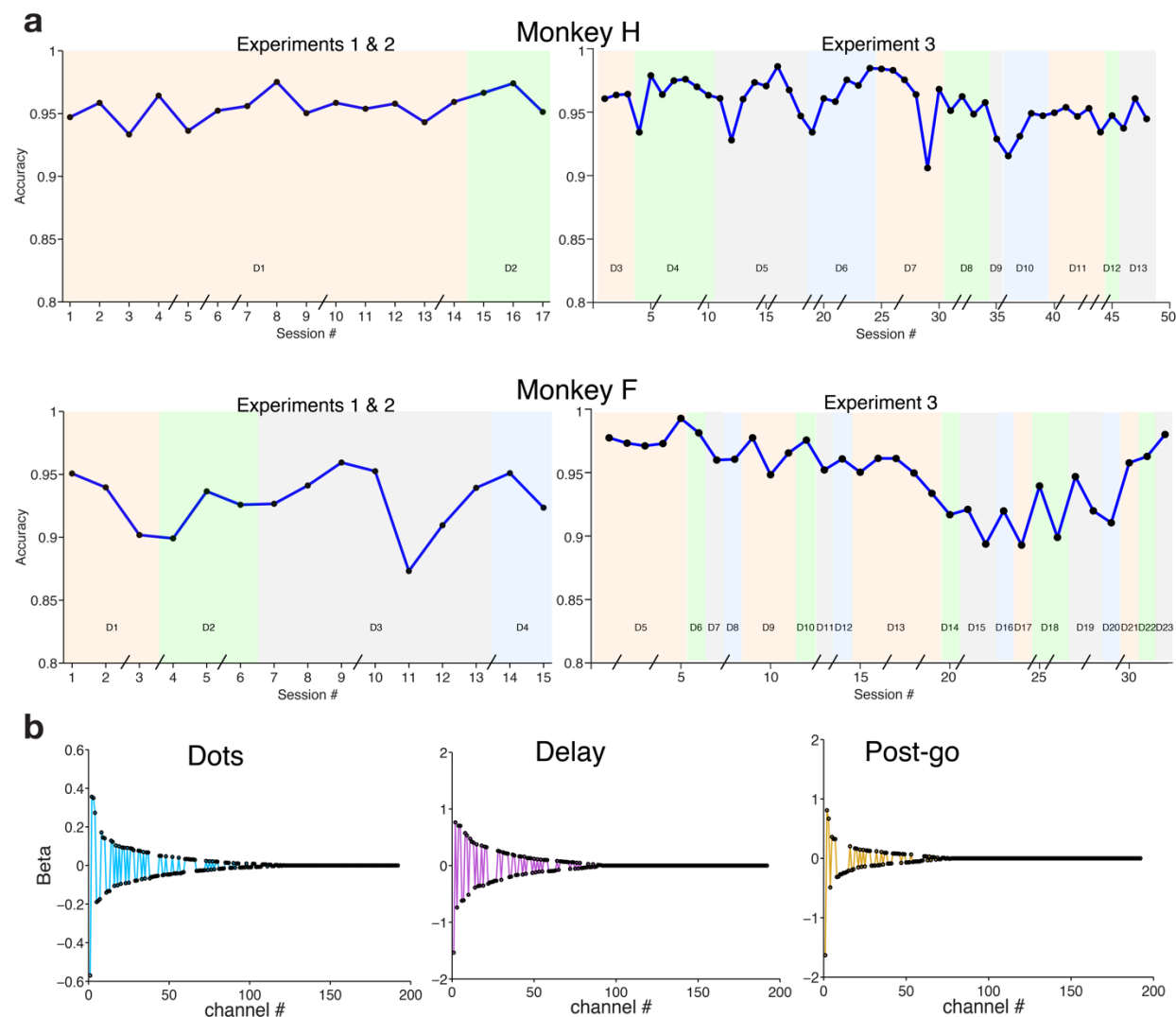
Extended Data Figure 5 – Real time decoding: performance reliability and decoder weights.

a) Decoding performance is stable across sessions. Average prediction accuracy during the second half of the stimulus presentation (600-1200 ms) across all sessions for monkey H (top panels) and monkey F (bottom panels). D1-D23 denote different decoders (sets of beta weights) used for the recorded sessions. For monkey H the same decoder (D1) was used for the first 14 sessions. The breaks on the x-axis correspond to sessions that occurred on non-consecutive days.

638   **b) Real time decoder beta weights.** Beta weights during the dots period (left panel) ranked by

639   absolute magnitude for an example decoder used in real time experiments. Channels with no or

640   little choice predictive activity during this period had their weights set to zero by using LASSO

641   regularization to prevent over fitting. Delay period and Post go cue Beta weights are shown in the

642   middle and right panels respectively.

643

644

645
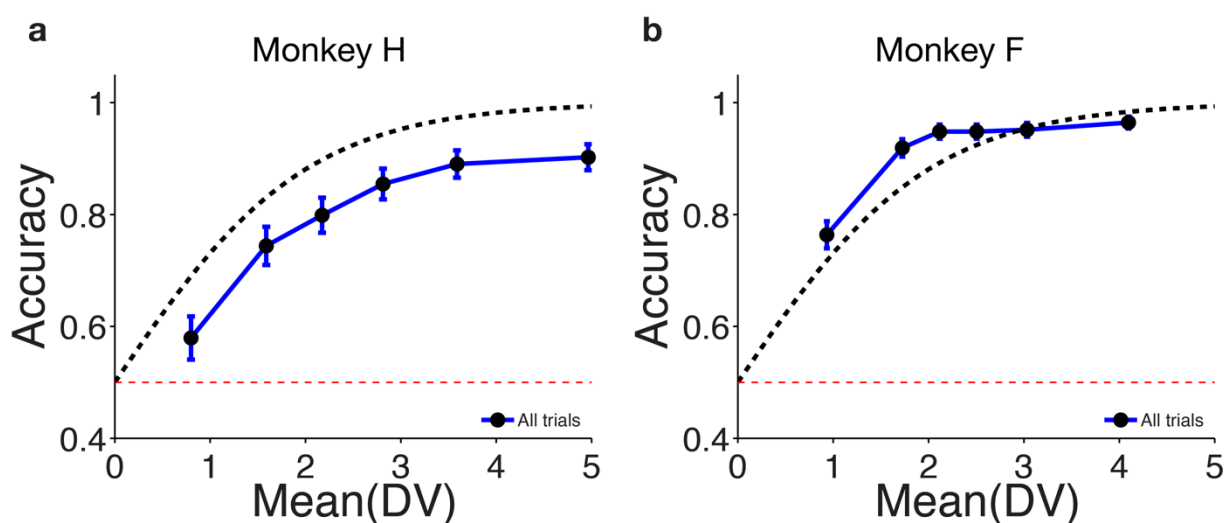
646   **Extended Data Figure 6 – Prediction accuracy as a function of DV for CoM trials.**

647

648   **a) Choice prediction accuracy for all trials collected during the CoM detection experiment -**

649   **monkey H.** Trials were split in 6 quantiles sorted by DV magnitude at termination. Prediction

650   accuracy and median DV magnitude was calculated and plotted separately for each quantile (blue

651   line with black markers). Blue error bars show standard error of the mean for a binomial

652   distribution. Dashed black line shows predicted accuracy from log-odds equation and red dashed

653   line shows chance level. Data from 985 CoM trials from monkey H.

654   **b) Choice prediction accuracy for all trials collected during the CoM detection experiment -**

655   **monkey F.** Same as **a)** for monkey F using 1727 CoM trials.

656

657
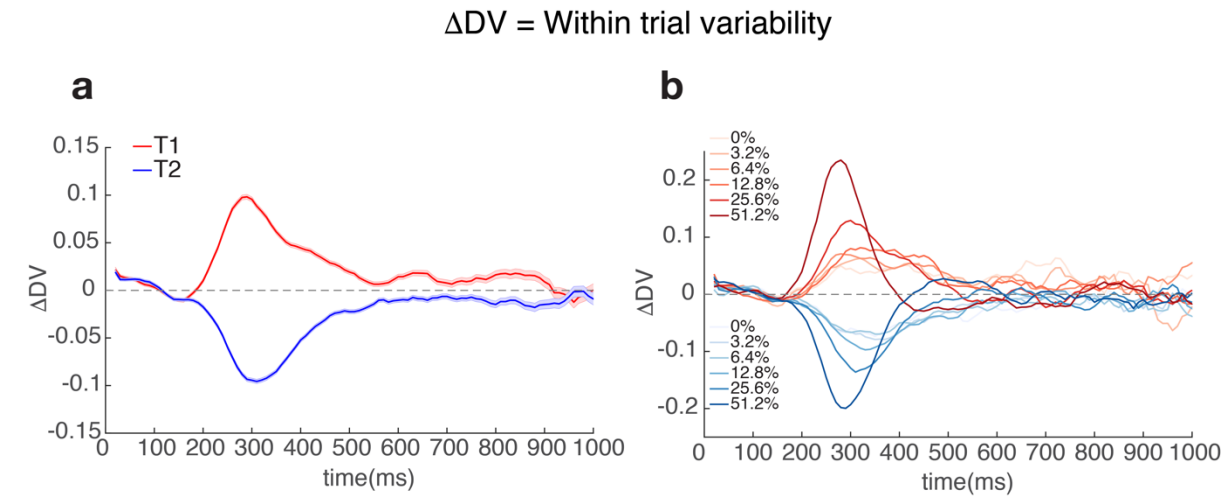
658   **Extended Data Figure 7 – Within trial variability as a function of time, choice and stimulus**

659   **coherence.**

660

661   **a) Average DV derivative as a function of time and choice - monkey F.** DV derivative was

662   calculated for each trial as the difference between consecutive DV estimates spaced out by 10 ms.

663   Traces show average DV derivative +/- s.e.m for right choices (red trace) and left choices (blue

664   trace).

665   **b) Average DV derivative as a function of time and signed coherence - monkey F.** Same data

666   as in **a)** but with DV derivative averaged separately for each choice and motion coherence level

667   (correct trials only). Right choices are plotted in red and left choices in blue as in **a)**. Darker traces

668   correspond to stronger coherences.

669

39
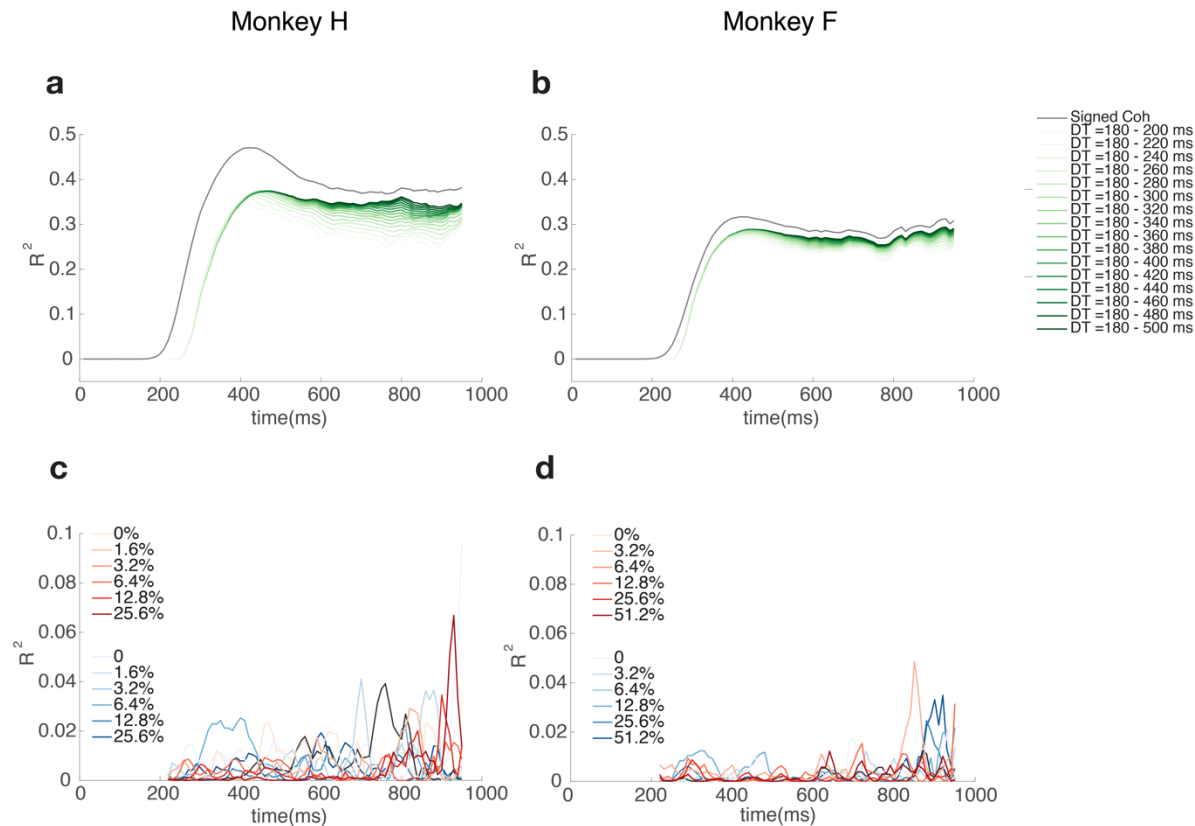
Extended Data Figure 8 – Correlation analysis between DV and stimulus motion energy.

a) Correlation between stimulus motion energy and decision variable - monkey H. Proportion of variance explained when regressing decision variable as a function of signed stimulus coherence (grey trace) or motion energy (green traces). Each green trace corresponds to a separate regression between DV and average motion energy between a timepoint in the past (from -200 ms up to -500 ms) and -180 ms (the estimated neural response delay). Darker traces correspond to regressions in which the motion energy was averaged for a longer time window. Across all coherence levels motion energy and signed coherence explain a large fraction of DV variance.

b Correlation between stimulus motion energy and decision variable - monkey F. Same as **a)** for monkey F.

682   **c) Correlation between stimulus motion energy and decision variable within each level of**

683   **signed coherence level - monkey H.** Proportion of variance explained when regressing DV for

684   each time point and within each level of signed coherence as a function of the motion energy

685   preceding it by 180 ms (the estimated neural response delay). Within each level of signed

686   coherence, the DV fluctuations are not predicted by the motion energy traces

687   **d) Correlation between stimulus motion energy and decision variable within each level of**

688   **signed coherence level - monkey F.** Same as **c)** for monkey F.

689

690   # Methods

691

692   **Subjects**

693

694   Our experiments were performed on two adult male macaque monkeys (*Macaca mulatta*) trained

695   to perform a direction discrimination task with reaching movements of the arm as operant

696   responses. These were the same subjects used in our previous study[10], but with new experiments.

697   All training, surgery, and recording procedures conformed to the National Institutes of Health

698   Guide for the Care and Use of Laboratory Animals and were approved by Stanford University

699   Animal Care and Use Committee.

700

701   **Apparatus**

702

703   Monkeys sat in a custom-made primate chair (Stanford Machine Shop) in front of a video

704   touchscreen, with their heads restrained using a surgical implant. The front plate of the chair could

705   be opened, allowing the subjects to reach the touchscreen with the arm contralateral to the

706   implanted hemisphere. The ispsilateral arm was gently restrained using a delrin tube and a cloth

707   sling. Stimuli were shown on the video touchscreen (ELO Touchsystems 1939L), which was

708   positioned approximately 35.5 cm away from the monkeys' heads and allowed hand position to be

709   tracked at 75Hz. Eye position was continuously tracked with an infrared eye tracker at 1kHz

710   (EyeLink 1000, SR Research, Canada).

711

712   **Motion Discrimination Task**

713

714   The task employed is a variation of the classical random dots motion discrimination task, in which

715   the subject uses an arm movement as the operant response[10] (Fig. 1a). We used a variable duration

716   version of this task in which the duration of the stimulus presentation varied from trial to trial.

717   There were two types of trials in our experiments: *open loop*, in which the stimulus duration was

718   determined by the experimenter at the beginning of the trial and *closed loop*, in which the duration

719   was contingent on a specific pattern of neural activity detected in real time (see Experiments 1-3).

720   The subject was never cued on what type of trial it was on. For open-loop trials stimulus duration

721   ranged from 500-1200 ms (median 670 ms) and was randomly chosen on each trial by sampling

722   an exponential distribution. For closed-loop trials the possible values for duration ranged between

723   250-1200 ms and were determined on each trial either by the timepoint at which the termination

724   conditions were met or a predetermined random duration sampled from the open loop distribution,

725   whichever came first. All trials started with the onset of a fixation point (FP; 1.5 degree diameter)

726   on the video touchscreen (Fig. 1a). To initiate the task, the monkey was required to maintain both

727   eye and hand fixation within +/- 3 degrees of the FP as long as it remained on the screen.

42

728   Importantly, throughout the entire trial, the monkey was required to always maintain direct hand

729   contact with the screen, otherwise the trial would be aborted.

730

731   After 300 ms of fixation, two targets (1.5 degree diameter) appeared on opposite sides of the FP

732   (eccentricities between 10 and 17 degrees). After a 500 ms delay the random dot stimulus was

733   presented for the durations mentioned above, after which it was removed from the screen. The

734   monkey was asked to report the net direction of motion (0 or 180 degrees) by reaching to the target

735   in the corresponding direction. The difficulty of the task was adjusted by changing the fraction of

736   dots moving coherently in one direction (motion strength). After stimulus offset the monkey either

737   entered a delay period during which it was required to withhold his response for 400-900 ms (on

738   30% of the open-loop trials) or was immediately presented the go cue (on 70% of the open-loop

739   trials and all closed-loop trials). The go cue was then signaled by the offset of the FP at which

740   point the monkey was free to gaze anywhere and report his decision with his arm by reaching one

741   of the two targets. Although gaze was monitored, reward acquisition depended solely on reaching

742   to the correct target. Finally, for a response to be considered valid, the monkey was required to

743   hold its hand position within +/- 4 degrees of the center of the target for 200 ms. The monkey was

744   then rewarded with a drop of juice for correct choices and given a timeout (2-4 seconds) for

745   incorrect ones. Zero coherence trials were rewarded randomly with a probability of 0.5 since there

746   was no correct response on these trials. The motion discrimination task was run on an Apple Mac

747   Pro running Mac OS.

748

749   **Random dots stimuli**

750

43

751   The stimuli used in our psychophysical experiment were random dot kinematograms (RDK)

752   generated using MATLAB and Psychophysics Toolbox. The details for generating the random

753   dots stimuli have been described previously[10]. However, to allow for closed loop experiments 1

754   and 2 (see below) we introduced a modification to be able to terminate the dot presentations early

755   if needed. The stimulus code was designed to precompute a sequence of kinematograms that

756   contain both random and moving dots. The sequence was then presented ballistically with no need

757   to continuously compute the content of each frame. Our modification allowed for DV values to be

758   received asynchronously from the real-time decoder and evaluated during the dots presentation. If

759   the DV criteria defined by the particular experiment were met, the dot presentation could then be

760   terminated without the remaining frames being shown. For the experiment in which an additional

761   pulse of motion energy was injected (closed loop experiment 3, see below), we arranged for two

762   sequences of kinematograms to be precomputed before presentation: one without the pulse, the

763   other for the 200 ms pulse itself. Contingent on the evolution of DV values, the stimulus could

764   then be rapidly switched from the standard sequence to the pulse sequence.

765

766   For both monkeys, the motion strength could take one of 6 possible values within a set, but the

767   sets were slightly different between subjects: [0%, 1.6%, 3.2%, 6.4%, 12.8%, 25.6%] for monkey

768   H and [0%, 3.2%, 6.4%, 12.8%, 25.6%, 51.2%] for monkey F. The top coherence (51.2%) was

769   dropped and a very low coherence (1.6%) was introduced for monkey H, due to its superior

770   discrimination ability. The direction and coherence of the motion were randomly assigned on each

771   trial by sampling from a uniform distribution with replacement. For zero-coherence stimuli all dots

772   were displaced randomly but, due to the stochasticity of that process, one obtains non-zero net

773   motion toward the targets over a small number of frames.

44

774

Behavioral Training

776

777   Both monkeys had been extensively trained on fixed and variable duration versions of the motion

778   discrimination task using an arm reach movement as the operant response prior to the current

779   study[10]. A few training sessions (all open-loop trials) were used to get the subject accustomed to

780   the new task timing (0.5-1.2 s stimuli and no delay on 70% of the trials). Real time decoding

781   sessions only started when psychophysical performance was stable.

782

Behavioral Analysis

784

785   Psychophysical performance was assessed in two ways: by describing the percentage of correct

786   choices as a function of (unsigned) stimulus coherence and by describing the percentage of

787   rightward choices as a function of signed stimulus coherence.

788

789   The percentage of correct choices as a function of motion strength (stimulus coherence) was fit by

790   a cumulative Weibull distribution function:

791

$$P_{correct}\ (c) = \ 1 - 0.5 \times\ e^{(-\frac{c}{\alpha})^{\beta}}$$

793

794   where $P_{correct}$ is probability correct, c is motion strength, α is the psychophysical threshold (the

795   value of c that corresponds to ~82% correct responses), and β is a parameter that controls the shape

796   of the function, especially its steepness.

797

798   The proportion of rightward choices, P_{right}, as a function of motion strength and direction was fit

799   by a logistic regression:

800

801    $$P_{right}(c) = \frac{1}{1 + e^{-\beta_1 \times (\beta_0 + c)}}$$

802

803   where c is motion strength, β_1 is the slope parameter and $-\beta_0$ is the motion strength corresponding

804   to the indifference point. This value was used to assess the monkey's behavioral bias on each

805   session.

806

807   **Electrophysiological recordings**

808

809   Two multielectrode arrays (Blackrock Microsystems, Utah) with 96 electrodes each (1mm long

810   platinum-iridium electrodes, 0.4 mm spacing, impedance average of approximately 400 KOhm)

811   were implanted in primary motor and dorsal premotor cortex of each monkey (Figure 1c). The

812   methods for determining the array placement were described in our previous study[10]. For monkey

813   F, the M1 array became unusable between the end of the previous study and the start of the current

814   study. Due to lack of neural signal from the M1 array, only the PMd array was used for this animal.

815   Continuous neural data were acquired and saved to disk from each channel (sampling rate 30 kHz)

816   and thresholded at -4.5 RMS using the Cerebus recording system (Blackrock Microsystems, Utah)

817   and two separate PCs (one for each array) running Windows 8. Waveforms corresponding to

818   threshold crossings were not sorted and each channel could contain one or more unit(s). Sorting

819   waveforms would require a significant lead-up time before the beginning of the experiment and

820   could negatively affect the ability to combine data and use decoders across days (see below,

821   Decoder training). Since units were not isolated within each channel our resulting units were most

822   likely multi-unit clusters. Any extremely noisy channels were deactivated at the beginning of a

823   session, and all other channels were used in this study. Using only multi-units yielded comparable

824   prediction accuracy (Extended Data Figure 3) to our previous study[10] in which both single and

825   multi-unit data was used.

826

827   **Datasets**

828

829   Data were collected in two sets of experiments. In the first set of experiments we performed Closed

830   Loop Experiments 1 and 2 (see below). For this set, for each monkey we analyzed all datasets that

831   met two behavioral inclusion criteria: 1) over 500 trials and 2) a behavioral bias ($|\beta_0|$) under 4%,

832   as determined by a logistic regression fit (see above). These criteria were imposed to ensure that

833   we have a sizeable number of trials per condition (6 coherence x 2 directions = 12 conditions) and

834   that the behavior of the monkey is virtually unbiased, such that both neural and behavioral results

835   are more easily interpretable. These criteria resulted in a selection of 17/15 sessions for a total of

836   16468/15826 trials for monkey H/F, respectively.

837

838   In the second set of experiments we performed Closed Loop Experiment 3. This set of experiments

839   was performed later, on separate sessions, but using the same two subjects, arrays and decoding

840   techniques as the first set. In this set of experiments, we analyzed all datasets with over 550 trials.

841   For all experiments in monkey H, PMD and M1 were recorded simultaneously.

842

843   **Decoder training**

844

845   We chose to use a logistic regression classifier based on our previous results showing excellent

846   offline prediction accuracy in variable duration tasks[10] and because of the direct probabilistic

847   interpretation of its output. Our decision variable (DV) was defined as the log odds ratio of

848   observing a particular behavioral choice (T_1 or T_2) given the population response $\vec{r}$:

849
$$DV = log\ \frac{P(T_1|\vec{r}\,)}{P(T_2|\vec{r}\,)} = \beta_0\,(t) + \sum_{i=1}^{n} \beta_i(t) \times r_i(t)$$

850

851   Where $r_i(t)$ are the z-scored summed spike counts for each neuron and time window, β_0 is an

852   intercept term and $\beta_i(t)$ are the classifier weights (one for each unit and epoch). Data from all

853   electrodes with valid waveforms were combined.

854

855   For simplicity, we decided to use only 3 different decoders for an entire trial (Fig. 1a), instead of

856   a different one for each 50 ms time window in the trial[10]. We applied the first decoder from fixation

857   up to and including the dots period, the second for the delay period and the third for the post go

858   cue period. After extensive offline tests on a few sessions the precise epochs for classifier training

859   were defined as the following:

860

861   • Dots epoch: [150, 1000] ms aligned to dots onset;

862   • Delay epoch: [250, 350] ms aligned to dots offset;

863   • Post-go cue epoch: [200,400] ms aligned to go cue;

864

865   LASSO regularization was applied to prevent over-fitting when calculating each set of beta

48

866   weights. A Lambda parameter constraining the L1 norm of the Beta vectors was calculated

867   separately for each of the 3 decoders using 10-fold cross validation on the corresponding time

868   epochs listed above. For each decoder the Lambda value with minimum cross-validation error was

869   chosen. Extended Data Figure 5b shows beta weights for an example set of 3 decoders for monkey

870   H sorted by epoch and ranked by magnitude. Positive weights correspond to rightward preferring

871   channels while negative weights correspond to leftward preferring channels. LASSO

872   regularization sets weights of channels with little or no predictive activity to zero.

873

874   The linear classifier was determined offline using recently collected data (from real-time

875   experiments). All 50 ms samples of neural data during the selected period (above) for each epoch

876   were used to train the classifier. The classifier was trained on 90% of the trials and tested on 10%

877   of the trials using 10-fold cross-validation. The weights from one of the cross-validation folds were

878   then used in the upcoming real-time experiments. Decisions to train new decoders were based on

879   experimenter judgment in attempts to optimize performance: if a substantial decrease in real-time

880   decoding performance and/or an increase in the DV offset at baseline was observed, a new

881   classifier was trained and used in the following session. New classifiers were typically used every

882   5 sessions, but some proved to be stable over up to 14 sessions (Extended Data Figure 5a).

883

884   **Real time decoding**

885

886   An essential requirement to compute a real-time read-out of neural activity is the ability to

887   continuously and (nearly) instantaneously access and perform computations on the neural activity

888   being recorded. To accomplish this, the spikes for each channel were temporally smoothed using

49

889   a causal half-Gaussian kernel with 50 ms standard deviation (to mitigate spurious Poisson

890   fluctuations) and summed for the most recent 50 ms. These smooth spike counts were then stored

891   in a 192x1 (96x1 for monkey F) vector of neural activity and z-scored individually for each

892   channel, using previously calculated μ (mean) and σ (standard deviation) vectors. Z-scoring neural

893   activity was crucial to ensure a reliable and stable real-time readout by preventing the highest firing

894   channels from dominating it. Finally, the z-scored neural activity was projected onto a previously

895   calculated linear decoder (a set of β weights, one for each channel) to obtain our linear readout of

896   internal decision state: a real time decision variable (DV)[1].

897

898   The value of the DV was updated every 10 ms, reflecting the neural activity of the preceding 50

899   ms. Because we used a half-gaussian kernel, data preceding the 50 ms window also influenced our

900   DV estimate (with more recent spikes carrying more weight). 95% of the data contributing to the

901   spike counts was limited to the last 100 ms (i.e an additional 50 ms in the past to each 50 ms

902   window). The DV value and its history on a single trial could then be used (if desired) to impose

903   conditions for termination of the random dots stimulus (experiments 1 and 2) or presentation of a

904   motion pulse (experiment 3), effectively closing the loop on the experiment.

905

906   While the β weights were not updated online (during the course of one experiment), the μ and σ

907   vectors for each epoch were learned continuously during the course of the experiment, due to

908   changing recording conditions and signals from day to day. The μ and σ vectors were initialized

909   at the beginning of the session using the values calculated offline when training the most recent

910   decoder. Once the session started, the initial μ and σ vectors were blended with online calculated

911   values for the first 25 trials, using a blending factor α:

912

913
$$\alpha_j = \ max\ ((25 - j)/25{,}0))\ , \text{ where } j \text{ is the trial number.}$$

914

915 For trial j, sample number t and for a given epoch in trial, the μ and σ vectors were defined as a

916 weighted mixture between the initial values $\mu_{initial}$(epoch) and $\sigma_{initial}$(epoch) and the estimate of the

917 current session's values $\mu_{current}$(t,epoch) and $\sigma_{current}$(t, epoch):

918
$$\mu_{blended}(t, epoch) = \ \alpha_j * \mu_{initial}(epoch) + (1 - \ \alpha_j) * \mu_{current}(t, epoch)$$

919
$$\sigma_{blended}(t, epoch) = \ \alpha_j * \sigma_{initial}(epoch) + (1 - \ \alpha_j) * \sigma_{current}(t, epoch)$$

920

921 After the first 25 trials α was set to zero which implies the μ and σ vectors kept being continuously

922 updated throughout the session but were no longer blended with values from the previous days.

923 The update rule for $\mu_{current}$(t, epoch) was:

924
$$\mu_{current}(t, epoch) = \frac{[\mu_{current}(t - 1, epoch)] * K + r}{K + 1}$$

925
$$K = \ N_{samples}(t, epoch)$$

926 where r is the most recently sampled vector of spike counts and K is the current number of samples

927 of spike count vectors obtained so far for this particular epoch.

928 The update rule for $\sigma_{current}$(t, epoch) was:

929
$$\sigma_{current}(t, epoch) = = \sqrt{\frac{K - 1}{K} * \sigma^2{}_{current}(t - 1, epoch) + \frac{1}{K} * (r - \mu_{current}(t, epoch))^2}$$

930 After updating the $\mu_{current}$(t, epoch) and $\sigma_{current}$(t, epoch) vectors, the number of

931 samples for the corresponding epoch was also updated:

932

933
$$N_{samples}(t, epoch) = N_{samples}(t - 1, epoch) + 1$$

934

935   Importantly, even though we had only 3 different decoders (Fig. 1b) we effectively used 5 different

936   epochs: Fixation, Targets, Dots, Delay and Post Go-Cue. The Dots decoder was also used in the

937   Fixation and Targets epochs but because average firing rates are different between these, different

938   μ and σ vectors had to be used. Every 50 ms sample of neural data for a given epoch was used to

939   update the corresponding μ and σ vectors as described above. We let the μ and σ vectors converge

940   for ~200-300 trials, in the beginning of each experimental session, before starting any closed loop

941   experiments. One way to check for this convergence was to monitor the DV offset: the average

942   DV value for the first 150 ms of the Dots epoch. Since we verified through offline analyses that

943   no systematic pre-planning activity towards one of the two targets was present in PMd or M1

944   during this time window, we expected the DV offset to be on average ~0.

945

946   Using a single decoder for an entire epoch was far more efficient to implement than using a

947   different decoder for each time point (as it reduced the number of μ and σ vectors that had to be

948   learned online) and as demonstrated in our previous study[10] a single classifier for an entire epoch

949   was almost as predictive as multiple classifiers trained on different timepoints of the same epoch[10].

950   Because choice modulation in PMd/M1 changes dramatically around the peri-movement period a

951   single decoder for an entire trial was not feasible.

952

953   In the end, our method yielded a reliable real-time decision state read out and required only ~18%

954   (15%) of trials in a session for calculating the values of μ and σ for monkey H (F), leaving the

955   remainder available for imposing neurally contingent conditions in closed loop. The real time

956 decoder was run on two separate PCs (server and client) using the Simulink Real-Time/xPC

957 platform (Mathworks, Massachussetts).

958

959 **Closed Loop Experiments**

960

961 *Experiment 1: Virtual Boundaries*

962

963 On each trial we set a virtual threshold, or boundary (B), for the magnitude of the DV during the

964 dots epoch. If the DV on the current trial reached B or −B ± tolerance, the dots presentation was

965 terminated and the monkey asked to report its decision. If the bound was not reached on a given

966 trial, stimulus presentation continued to a preset duration for that trial which was randomly

967 sampled from an exponential distribution ranging from 500-1200 ms. Closed loop trials for which

968 the boundary was not reached were effectively indistinguishable from open loop trials.

969

970 Typically, 5 values for boundaries spanning 0.5 to 5 (DV units) were used every session and one

971 of them was randomly assigned on each trial (uniform distribution). The tolerance used was ±

972 0.25 DV units. We imposed a minimum duration for all trials to avoid spurious bound crossings,

973 which could be problematic for low bound values in particular. In all sessions the minimum

974 duration was 250 ms, a conservative estimate of the latency for choice related signals driven by

975 the visual stimulus to appear in PMd and M1.

976

977 After the minimum stimulus duration was reached, the DV was assessed every 10 ms to determine

978 whether it fell within ± 0.25 DV units of the bound chosen for the current trial (B or −B). If so, the

53

979   stimulus was terminated within 34 ms of the boundary being met (see Estimated latency for real

980   time closed loop setup). If the bound for the particular trial was not reached, the presentation

981   continued up to the maximum stimulus duration selected for that trial which had been obtained by

982   randomly sampling from an exponential distribution: 500-1200 ms (median 670 ms).

983

984   Finally, closed loop trials were randomly interleaved with open loop trials in which no DV-

985   dependent termination condition was imposed. The motivation for interleaving closed loop and

986   open loop trials was to make it extremely hard for the monkey to learn that accelerating the

987   dynamics of choice related signals[10] (potentially by recruiting more choice related neurons or

988   increasing their modulation) and thus hitting bounds sooner could potentially increase its reward

989   rate. Not accounting for this possibility could lead to an undesirable change in the monkey's

990   strategy during the course of the closed loop experiments, which could become problematic when

991   combining data across days.

992

993   *Experiment 2: CoM detection*

994

995   Under our logistic regression framework, the signature of a putative CoM is a sign change of the

996   decision variable. Since these sign changes could happen at any time during the trial, capturing

997   them required not only monitoring the most recent state of the DV, but its history throughout the

998   trial. Because there was noise in our DV estimation and DVs usually started close to 0 at the

999   beginning of the trial we imposed selection criteria to detect likely CoMs based on the neural data.

1000   A necessary feature for all potential CoMs was a zero crossing in the sign of the DV: change of

1001   DV sign from negative to positive reflected a change in the likelihood of a rightward decision from

54

1002   less than 50% to greater than 50%, and vice versa for the opposite change in sign. To eliminate

1003   zero crossings resulting solely from measurement noise, we imposed four additional criteria:

1004

1005   • Minimum DV value after zero crossing;

1006   • Minimum DV value with opposite sign before zero crossing;

1007   • Minimum duration of DV sign stability after zero crossing;

1008   • Minimum duration of DV sign stability before zero crossing;

1009

1010   The minimum DV values before and after zero crossing were symmetrical for most sessions, as

1011   were the periods of minimum duration of DV sign stability (negative or positive values for all time

1012   points). If a zero crossing was detected and all four criteria were met, the stimulus presentation

1013   was interrupted and the animal was virtually immediately (within 34 ms or less, see Estimated

1014   latency for real time closed loop setup) prompted to report a decision. The exact parameters used

1015   for each session can be found in Supplementary Information Table 3.

1016

1017   By sweeping the parameter space we could test zero crossings that differed in magnitude and

1018   stability. Analogously to the virtual boundary experiment, if the minimums were not met and a

1019   CoM thus not detected, the stimulus presentation continued uninterrupted for a random duration

1020   ranging from 500-1200 ms, selected prior to the start of the trial. A minimum stimulus duration of

1021   250 ms was also in place.

1022

1023   Because putative CoMs are quite rare[1], in the first set of experiments we devoted 70% of the closed

1024   loop trials to detect them leaving the remaining 30% as virtual boundary trials. The exact fraction

1025   of trials with CoM depends dramatically on how we parameterize them. The longer the minimum

1026   periods of consistent sign and the higher the minimum DV value in the initial commitment stage,

1027   the rarer they become. Running both experiments on the same sessions ensured that the mapping

1028   from DV to choice likelihood was held during the CoM experiments and provided the most faithful

1029   indirect validation of initial commitment we could obtain.

1030

1031   *Experiment 3: Motion pulse perturbation*

1032

1033   In this experiment, motion pulses were introduced on some trials with motion coherences near or

1034   below psychophysical threshold. No motion pulses were presented for suprathreshold coherences

1035   based on the results of a pilot experiment (not shown) in which pulses presented at suprathreshold

1036   coherences were more perceptually salient and led to changes in the animals' strategy. As in

1037   Experiment 1, on each trial we set a virtual boundary (B) for the magnitude of the DV during the

1038   dots epoch. In this experiment, 100% of trials with dots coherence at or near psychophysical

1039   threshold were treated as closed loop trials (this corresponds to trials with maximum unsigned

1040   coherence of 6.4% for monkey H and 12.8% for monkey F; psychophysical thresholds were

1041   measured using the Weibull function described above in "Behavioral Analysis"). Low-coherence

1042   trials in which the boundary was not reached (per the criteria below) and trials with suprathreshold

1043   dots coherences were all effectively open loop.

1044

1045   If the DV on a closed loop trial reached B or $-B \pm$ tolerance (±0.25 DV units), after a minimum

1046   stimulus duration of 50 ms, a 200-ms motion pulse was presented, followed immediately by

1047   termination of the visual stimulus and presentation of the cue for the monkey to report its decision.

1048   If not, dots presentation continued for a pre-assigned duration drawn randomly from an

1049   exponential distribution of 500-1200 ms. Four integer values for boundaries (spanning 1 to 4 DV

1050   units) were used every session, and one of them was randomly assigned on each trial (uniform

1051   distribution).

1052

1053   Motion pulses were 200-ms periods of additional dots stimulus presentation with small additive

1054   average coherence (±2% or 4.5% from the initial dots coherence on the same trial for monkey H

1055   and F, respectively, where positive coherence values indicate rightward motion); thus pulses

1056   effectively randomly added either a small amount of rightward or leftward motion evidence to the

1057   stimulus. Pulse strength was calibrated in pilot experiments, in which we converged upon

1058   coherence shifts that slightly but significantly biased each animal's behavior, without being overtly

1059   perceptually salient (biases were measured using the logistic regression on rightward choice

1060   described above in "Behavioral Analysis"). Animals were rewarded for correct reaches in the

1061   direction of the coherence of the initial dots stimulus (randomly assigned on 0% coherence trials),

1062   regardless of the pulse direction.

1063

1064   *Estimated latency for real time closed loop setup*

1065

1066   To validate our setup, we measured the latency between a neural condition being met and the

1067   corresponding task change being implemented. We tested this latency by generating simulated DV

1068   steps in the same model used to detect when DV triggering conditions were met in the real

1069   experiments. We used these simulated steps to trigger the onset of a bright light on the touchscreen

1070   in front of a photodetector, again within the same code used to run the task and generate the stimuli

1071 in the real experiments. We then passed both the simulated DV and the photodetector output

1072 signals into an oscilloscope, triggered the display on the "DV" steps, and manually measured the

1073 delay to onset of the bright dot. Almost all measured delays were within 2 frames, or 26 ms.

1074

1075 *Estimated trial count savings for real time closed loop setup*

1076

1077 The real time setup allowed for extremely precise experimental control over which DV values or

1078 DV history to use to trigger a modification in the task (stimulus termination or pulse). However, it

1079 could be argued that given enough data, similar trials would have been captured simply by either

1080 terminating the stimulus (as in experiments 1 and 2) or presenting the pulse (as in experiment 3)

1081 at a random point in the trial and then back sorting them offline (by DV value or history by after

1082 the data is collected).

1083

1084 To estimate how much more trial-count efficient it was to use our real time setup compared to

1085 offline back-sorting trials where the stimulus was presented for a random duration, we used the

1086 CoM experiment as a case study given how rare change of mind events are.

1087

1088 For simplicity, we focused on sessions 1, 2 and 3 from Monkey F, which all have the same (and

1089 intermediate) CoM requirements (Supplementary Information Table 3). We started by calculating

1090 the yield from the real time experiment in closed loop as the ratio between detected CoM trials and

1091 trials in which CoMs were checked (i.e. all closed loop trials in which the stimulus could be

1092 terminated if the conditions dictated by the CoM parameters were met, Supplementary Information

1093 Table 3):

58

1094 $$Yield_{CL} = \#CoMs\ detected/\#trials\ CoMs\ checked = 11.91\%$$

1095

1096 To calculate the yield for offline back-sorting trials we used the open loop trials in the same

1097 sessions, which were terminated after a random stimulus duration. Importantly, the stimulus

1098 duration on these open loop trials was sampled from the same distribution as for the closed loop

1099 trials in which CoMs were checked, which allows for a fair yield comparison. We calculated the

1100 yield from offline back-sorting as the ratio between the number of trials that would have met all

1101 the criteria for CoMs for the same session and the total number of open loop trials:

1102

1103 $$Yield_{OL} = \#Valid\ putative\ CoMs\ /\#Open\ loop\ trials = 1.85\%$$

1104

1105 Since the goal would be to probe the new choice preference shortly after the zero crossing (putative

1106 change of mind), not many hundreds of ms later, we only considered CoM trials that were (closed

1107 loop) or that would have been (open loop) terminated within 150 ms of the zero crossing. This

1108 cutoff value corresponded to the 82nd percentile of post zero crossing durations for the closed loop

1109 trials analysed in these sessions.

1110

1111 In this analysis $Yield_{CL}$ was 6.43 times higher than $Yield_{OL}$. This result implies that had we not

1112 used a real time setup in closed loop we would have had to collect 6.43 times the number of trials

1113 (and thus sessions) to obtain the same number of events. This would in turn mean collecting

1114 around 100 sessions/monkey just for experiments 1 and 2 (assuming the same 30%/70% trial split

1115 used in the real time experiments), rendering this experiment practically unfeasible.

1116

1117 **Neural Data Analysis**

1118

1119 *DV variability*

1120

1121 Within trial variability was computed by first calculating the difference between consecutive DV

1122 values (estimated every 10 ms) for every trial in the datasets collected for experiments 1 and 2

1123 (open and closed loop). This step yielded a ΔDV trace for each trial aligned to dots onset. For each

1124 trial these traces were computed only up to the offset of the stimulus and did not include any delay

1125 or post go-cue DV data. The ΔDV traces were then sorted and averaged for each choice (Fig. 5a,

1126 Extended Data Fig.7a) or each signed coherence level (Fig. 5b, Extended Data Fig.7b). Longer

1127 trials are increasingly rare due to the shape of our stimulus duration distribution but this asymmetry

1128 does not influence the interpretation of the time course of average ΔDV as this metric only captures

1129 *within* trial variability and not *across* trial variance.

1130

1131 *DV and motion energy correlation*

1132

1133 Motion energy (ME) was calculated for each trial in the datasets collected for experiments 1 and

1134 2 (open and closed loop) by convolving the positions of the dots in the stimulus with spatio-

1135 temporal filters as previously described[8]. The ME trace obtained for each trial captures the strength

1136 of the stimulus at every timepoint during the stimulus presentation. To evaluate the effect of motion

1137 energy on DV we performed a linear regression of single trial DV traces on single trial ME traces.

1138 From experiments 1 and 3 we determined that due to neural latencies a stimulus fluctuation can

1139 only have effect on the decoded DV ~180 ms later. For this reason, the regression was always

60

1140 performed between DV(t) and ME(t-180ms) or earlier. For Extended Data Fig. 8a-b each green

1141 trace corresponds to a different way to estimate the motion energy that might affect DV at time t.

1142 For the lightest trace and for every timepoint t ME was averaged between (t-180ms) and (t-200ms)

1143 for every trial and used to regress against DV(t). A separate regression was performed for each

1144 timepoint and the resulting variance explained was plotted. The same process then was repeated

1145 for every other green trace by progressively increasing the averaging window for ME in 20 ms

1146 increments from (t-180, t- 200) ms to (t-180, t- 500) ms. As a control the DV was also regressed

1147 against signed coherence for each trial (Extended Data Fig. 8a-b grey traces). This analysis was

1148 used to assess how much of the DV variance *across* coherences is explained by motion energy or

1149 signed coherence as a function of time. To assess how much DV variance *within* each coherence

1150 level could be explained by the motion energy of the stimulus we first sorted the DV traces for

1151 each signed coherence level. For each signed coherence level and each timepoint we regressed

1152 DV(t) against ME(t-180ms) for the corresponding trials and calculated the variance explained

1153 (Extended Data Fig. 8c-d).

1154

1155 *CoM regularities*

1156

1157 To test whether the effects of coherence on the number of CoMs were statistically significant we

1158 used a bootstrap method to generate 1000 distributions of CoM events with the corresponding

1159 coherences by sampling with replacement from the distribution of captured events for each subject

1160 separately. For each subject each distribution had the same number of observations as those

1161 captured in experiment 2: 985 for monkey H and 1727 for monkey F. For each randomly sampled

1162 distribution the number of CoMs for each coherence level was counted. The resulting counts were

1163   then regressed against the coherence level they belonged to. CoM count was highly and negatively

1164   correlated with coherence for both subjects (p<0.001).

1165   To test whether the effect of directionality of CoMs (corrective vs erroneous) was statistically

1166   significant we followed a similar bootstrapping procedure and generated 1000 distributions of

1167   CoM events (excluding 0% trials) with the corresponding directionality. For each randomly

1168   sampled distribution the number of CoMs for each direction was counted. The difference between

1169   the medians counts of corrective and erroneous CoMs was tested by performing a one-sided

1170   Wilcoxon rank sum test (p<0.001) testing the hypothesis than corrective counts were higher than

1171   erroneous counts.

1172

1173   *Pulse Effects*

1174

1175   To quantify the overall behavioral effect of the pulses, we performed the following logistic

1176   regression on the probability of a rightward choice:

1177

1178
$$P_{right}(c) = \frac{1}{1 + e^{-\beta_1 \times \left(\beta_0 + \beta_{pulse}D + c \right)}},$$

1179

1180   where c is motion strength, β_1 is the slope parameter, D is the pulse direction, and $-\beta_0$ is the motion

1181   strength corresponding to the indifference point.

1182

1183   To determine the effect of the pulse on the evolving DV, we first estimated the minimum latency

1184   for the visual stimulus information to influence the DV by calculating the first time of significant

1185   divergence of rightward vs. leftward DV traces during dots presentation on open-loop trials with

1186   stimuli of maximal motion strength (±25.6%, 51.2% coherence for monkey H, F), assessed using

1187   a two-sample t-test with correction for a false discovery rate of 0.05[34]. We refer to this estimate as

1188   the "evidence representation latency" (ERL; 170 ms, 180 ms for monkey H, F). We then measured

1189   the evolution of the DV after pulse presentation by calculating the difference between the

1190   empirically observed DV at each time point t and the DV at the "pulse evidence representation

1191   latency" (PERL, or time of pulse onset plus the ERL):

1192

1193
$$\Delta DV(t) = DV(t) - DV_{PERL}.$$

1194

1195   To quantify the pulse effect on DV on single trials, we calculated the slope (using linear regression)

1196   of ΔDV over the time period beginning with the PERL and ending at either the median go-RT for

1197   the animal or 150 ms prior to movement initiation on that trial, whichever came first, to minimize

1198   the contribution of directly movement-related activity to our analysis of the evolving choice[10]. We

1199   checked to ensure the results did not depend critically on these specific endpoint criteria by

1200   sweeping a range of endpoints for the ΔDV calculation: every 10 ms from go cue onset to 150 ms

1201   after the go cue. The exact endpoint used did not affect the results of the following analyses

1202   (quantitative results held for all endpoints from the time of the go cue through 150 ms post-go-cue

1203   for monkey H, and for endpoints beginning 120 ms post-go-cue for monkey F).

1204

1205   To quantify the neural pulse effect at each DV boundary (as in Fig. 5e), we fit the following linear

1206   regression on the ΔDV slope (calculated as described in the previous paragraph) separately for

1207   each DV boundary:

1208

1209
$$\Delta DV\ slope^B = \text{ß}_0^B + \text{ß}_{coh}^B \times c + \text{ß}_{time}^B \times time_{pulse\ onset} + \text{ß}_{pulse}^B \times D$$

1210
$$+\ \text{ß}_{pulse \times time}^B \times \left(time_{pulse\ onset} \times D\right),$$

1211

1212   where B is the pulse-triggering DV boundary (unsigned), and compared ß_{pulse}^B values. (The full

1213   resulting sets of regression coefficients fit to ΔDV slope can be found in Supplementary

1214   Information Table 1.)

1215

1216   Similarly, to quantify the behavioral pulse effect at each DV boundary (as in Fig. 5f), we fit the

1217   following logistic regression on the probability of a rightward choice separately for each DV

1218   boundary:

1219
$$P_{right}^B(c) = \frac{1}{1 + e^{-(z^B)}},$$

1220
$$\text{where } z^B = \text{ß}_0^B + \text{ß}_{coh}^B \times c + \text{ß}_{time}^B \times time_{pulse\ onset} + \text{ß}_{pulse}^B \times D$$

1221
$$+\ \text{ß}_{pulse \times time}^B \times \left(time_{pulse\ onset} * D\right),$$

1222

1223   where B is again the pulse-triggering DV boundary (unsigned), and compared ß_{pulse}^B values. (The

1224   full resulting sets of regression coefficients fit to choice can be found in Supplementary

1225   Information Table 2.)

1226

1227   **Data availability**

1228

1229   All data and analyses generated during the current study are available from the corresponding

1230   authors upon reasonable request.

1231

1232 **Code availability**

1233

1234 The analysis code was developed in Matlab (Mathworks) and is available from the corresponding

1235 authors upon reasonable request.

1236

1237

1238

1239

1240

References

1242

1243 1 Kiani, R., Cueva, C. J., Reppas, J. B. & Newsome, W. T. Dynamics of neural population

1244 responses in prefrontal cortex indicate changes of mind on single trials. *Curr Biol* **24**,

1245 1542-1547, doi:10.1016/j.cub.2014.05.049 (2014).

1246 2 Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Vacillation, indecision

1247 and hesitation in moment-by-moment decoding of monkey motor cortex. *Elife* **4**, e04677,

1248 doi:10.7554/eLife.04677 (2015).

1249 3 Bollimunta, A., Totten, D. & Ditterich, J. Neural Dynamics of Choice: Single-Trial

1250 Analysis of Decision-Related Activity in Parietal Cortex. *The Journal of Neuroscience*

1251 **32**, 12684 (2012).

1252   4   van den Berg, R. *et al.* A common mechanism underlies changes of mind about decisions

1253   and confidence. *eLife* **5**, e12192, doi:10.7554/eLife.12192 (2016).

1254   5   Lemus, L. *et al.* Neural correlates of a postponed decision report. *Proceedings of the*

1255   *National Academy of Sciences* **104**, 17174, doi:10.1073/pnas.0707961104 (2007).

1256   6   Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nat*

1257   *Neurosci* **19**, 973-980, doi:10.1038/nn.4320 (2016).

1258   7   Resulaj, A., Kiani, R., Wolpert, D. M. & Shadlen, M. N. Changes of mind in decision-

1259   making. *Nature* **461**, 263, doi:10.1038/nature08275

1260   https://www.nature.com/articles/nature08275#supplementary-information (2009).

1261   8   Kiani, R., Hanks, T. D. & Shadlen, M. N. Bounded integration in parietal cortex

1262   underlies decisions even when viewing duration is dictated by the environment. *J*

1263   *Neurosci* **28**, 3017-3029, doi:10.1523/JNEUROSCI.4761-07.2008 (2008).

1264   9   Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual

1265   motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**,

1266   4745-4765 (1992).

1267   10   Peixoto, D. *et al.* Population dynamics of choice representation in dorsal premotor and

1268   primary motor cortex. *bioRxiv* (2018).

1269   11   Shadlen, M. N. & Newsome, W. T. Neural Basis of a Perceptual Decision in the Parietal

1270   Cortex (Area LIP) of the Rhesus Monkey. *J Neurophysiol* **86**, 1916-1936 (2001).

1271   12   Ratcliff, R. & McKoon, G. The Diffusion Decision Model: Theory and Data for Two-

1272   Choice Decision Tasks. *Neural Computation* **20**, 873-922, doi:10.1162/neco.2008.12-06-

1273   420 (2007).

1274   13   Mazurek, M. E. A Role for Neural Integrators in Perceptual Decision Making. *Cerebral*

1275   *Cortex* **13**, 1257-1269, doi:10.1093/cercor/bhg097 (2003).

1276   14   Lo, C.-C. & Wang, X.-J. Cortico-basal ganglia circuit mechanism for a decision threshold

1277   in reaction time tasks. *Nat Neurosci* **9**, 956-963,

1278   doi:http://www.nature.com/neuro/journal/v9/n7/suppinfo/nn1722_S1.html (2006).

1279   15   Usher, M. & McClelland, J. L. The time course of perceptual choice: The leaky,

1280   competing accumulator model. *Psychological Review* **108**, 550-592, doi:10.1037/0033-

1281   295X.108.3.550 (2001).

1282   16   Cisek, P., Puskas, G. A. & El-Murr, S. Decisions in Changing Conditions: The Urgency-

1283   Gating Model. *The Journal of Neuroscience* **29**, 11560 (2009).

1284   17   Huk, A. C. & Shadlen, M. N. Neural Activity in Macaque Parietal Cortex Reflects

1285   Temporal Integration of Visual Motion Signals during Perceptual Decision Making. *The*

1286   *Journal of Neuroscience* **25**, 10420 (2005).

1287   18   Kiani, R. & Shadlen, M. N. Representation of confidence associated with a decision by

1288   neurons in the parietal cortex. *Science (New York, N.Y.)* **324**, 759-764,

1289   doi:10.1126/science.1169405 (2009).

1290   19   Smith, P. L. & Ratcliff, R. Psychology and neurobiology of simple decisions. *Trends in*

1291   *Neurosciences* **27**, 161-168, doi:10.1016/j.tins.2004.01.006 (2004).

1292   20   Ditterich, J. Evidence for time-variant decision making. *European Journal of*

1293   *Neuroscience* **24**, 3628-3641, doi:10.1111/j.1460-9568.2006.05221.x (2006).

1294   21   Hanks, T., Kiani, R. & Shadlen, M. N. A neural mechanism of speed-accuracy tradeoff in

1295   macaque area LIP. *eLife* **3**, e02260, doi:10.7554/eLife.02260 (2014).

1296   22   Wong, K.-F., Huk, A., Shadlen, M. & Wang, X.-J. Neural circuit dynamics underlying

1297   accumulation of time-varying evidence during perceptual decision making. *Frontiers in*

1298   *Computational Neuroscience* **1**, 6 (2007).

1299   23   Inagaki, H. K., Fontolan, L., Romani, S. & Svoboda, K. Discrete attractor dynamics

1300   underlies persistent activity in the frontal cortex. *Nature* **566**, 212-217,

1301   doi:10.1038/s41586-019-0919-7 (2019).

1302   24   Seidemann, E., Meilijson, I., Abeles, M., Bergman, H. & Vaadia, E. Simultaneously

1303   recorded single units in the frontal cortex go through sequences of discrete and stable

1304   states in monkeys performing a delayed localization task. *The Journal of Neuroscience*

1305   **16**, 752 (1996).

1306   25   Golub, M. D., Chase, S. M., Batista, A. P. & Yu, B. M. Brain–computer interfaces for

1307   dissecting cognitive processes underlying sensorimotor control. *Current Opinion in*

1308   *Neurobiology* **37**, 53-58, doi:https://doi.org/10.1016/j.conb.2015.12.005 (2016).

1309   26   Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Making brain–

1310   machine interfaces robust to future neural variability. *Nature Communications* **7**, 13749,

1311   doi:10.1038/ncomms13749

1312   https://www.nature.com/articles/ncomms13749#supplementary-information (2016).

1313   27   Semedo, J. D., Zandvakili, A., Machens, C. K., Yu, B. M. & Kohn, A. Cortical Areas

1314   Interact through a Communication Subspace. *Neuron* **102**, 249-259.e244,

1315   doi:https://doi.org/10.1016/j.neuron.2019.01.026 (2019).

1316   28   Seideman, J. A., Stanford, T. R. & Salinas, E. Saccade metrics reflect decision-making

1317   dynamics during urgent choices. *Nature Communications* **9**, 2907, doi:10.1038/s41467-

1318   018-05319-w (2018).

1319 29 Musall, S., Kaufman, M. T., Juavinett, A. L., Gluf, S. & Churchland, A. K. Single-trial

1320   neural dynamics are dominated by richly varied movements. *bioRxiv*, 308288,

1321   doi:10.1101/308288 (2019).

1322 30 Aflalo, T. *et al.* Decoding motor imagery from the posterior parietal cortex of a

1323   tetraplegic human. *Science* **348**, 906 (2015).

1324 31 Andersen, R. A., Hwang, E. J. & Mulliken, G. H. Cognitive Neural Prosthetics. *Annual*

1325   *Review of Psychology* **61**, 169-190, doi:10.1146/annurev.psych.093008.100503 (2009).

1326 32 Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive

1327   Control Signals for Neural Prosthetics. *Science* **305**, 258 (2004).

1328 33 Pesaran, B., Musallam, S. & Andersen, R. A. Cognitive neural prosthetics. *Current*

1329   *Biology* **16**, R77-R80, doi:10.1016/j.cub.2006.01.043 (2006).

1330 34 Benjamini, Y. & Hochberg, Y. *Controlling The False Discovery Rate - A Practical And*

1331   *Powerful Approach To Multiple Testing*. Vol. 57 (1995).

1332

1333 **End notes**

1334

1335 **Supplementary Information** is linked to the online version of the paper.

1336

1337 **Acknowledgments**

1338

1353

Author Contributions

1355

1356   D.P., J.R.V., R.K., S.F., K.V.S. and W.T.N. designed the experiments. D.P., J.R.V. and S.F. trained

1357   the animals and collected the data. D.P., J.R.V. and W.T.N. wrote initial draft of the paper. S.I.R,

1358   D.P. and R.K. performed the surgical procedures. D.P., J.C.K., P.N., C.C. and J.B. implemented

1359   the real-time decoding setup. D.P., R.K. and C.C. designed, and D.P. and J.R.V implemented, the

1360   decoder training algorithm to obtain the decoder weights and normalization matrices. D.P. and

1361   J.R.V. analysed the data. All authors contributed analytical insights and commented on statistical

1362   tests, discussed the results and implications, and contributed extensively to the multiple subsequent

1363   drafts of the paper.

1364

Author Information

1365

1366

1367 The authors declare no competing interests. Correspondence and requests for materials should be

1368 addressed to D.P. (dpeixoto@stanford.edu), J.R.V. (jverhein@stanford.edu), or W.T.N.

1369 (bnewsome@stanford.edu).

1370