

1 **Electrical coupling controls dimensionality and chaotic firing of**  
2 **inferior olive neurons**

3

4 Huu Hoang<sup>1,9\*</sup>, Eric J. Lang<sup>2</sup>, Yoshito Hirata<sup>3,4,5</sup>, Isao T. Tokuda<sup>6</sup>, Kazuyuki Aihara<sup>3,5</sup>, Keisuke  
5 Toyama<sup>1</sup>, Mitsuo Kawato<sup>1,7\*</sup> and Nicolas Schweighofer<sup>8\*</sup>

6

7 <sup>1</sup> Computational Neuroscience Laboratories, ATR Institute International, Kyoto, Japan

8 <sup>2</sup> Department of Neuroscience and Physiology, New York University School of Medicine, New  
9 York, USA

10 <sup>3</sup> Institute of Industrial Science, The University of Tokyo, Tokyo, Japan

11 <sup>4</sup> Mathematics and Informatics Center, The University of Tokyo, Tokyo, Japan

12 <sup>5</sup> International Research Center for Neurointelligence (WPI-IRCN), The University of Tokyo,  
13 Tokyo, Japan

14 <sup>6</sup> Department of Mechanical Engineering, Ritsumeikan University, Shiga, Japan

15 <sup>7</sup> RIKEN Center for Advanced Intelligence Project, ATR Institute International, Kyoto, Japan

16 <sup>8</sup> Biokinesiology and Physical Therapy, University of Southern California, California, USA

17 <sup>9</sup> Lead contact

18 \* *Corresponding authors: H.H (hoang@atr.jp), M.K (kawato@atr.jp) and N.S*  
19 *(schweigh@pt.usc.edu)*

20 **SUMMARY**

21

22 One of the main challenges facing online neural learning systems with numerous modifiable  
23 parameters (or “degrees-of-freedom”) such as the cerebellum, is how to avoid “overfitting”. We  
24 previously proposed that the cerebellum controls the degree-of-freedom during learning by  
25 gradually modulating the electric coupling strength between inferior olive neurons. Here, we  
26 develop a modeling technique to estimate effective coupling strengths between inferior olive  
27 neurons from *in vivo* recordings of Purkinje cell complex spike activity in three different coupling  
28 conditions. We show that high coupling strengths induce synchronous firing and decrease the  
29 dimensionality of inferior olive firing dynamics. In contrast, intermediate coupling strengths  
30 induce chaotic firing and increase the dimensionality of firing dynamics. Our results thus support  
31 the hypothesis that effective coupling controls the dimensionality of inferior olive firing, which  
32 may allow the olivocerebellar system to learn effectively from a small training sample set despite  
33 the low firing frequency of inferior olive neurons.

## 34 INTRODUCTION

35

36 Fifty years ago, David Marr proposed a theory that synaptic plasticity in the cerebellar cortex  
37 induces motor learning (Marr, 1969). Multiple experimental and computational studies have  
38 since supported this theory: climbing fiber afferents carry error signals (Bazzigaluppi et al.,  
39 2012; Keating and Thach, 1995; Kitazawa et al., 1998; Kobayashi et al., 1998), that modify  
40 parallel-fiber-Purkinje-cell synapses (D'Angelo et al., 2016; Hansel et al., 2001; Ito, 2001;  
41 Kuroda et al., 2001), and drive learning of internal models for motor control (Bastian, 2006;  
42 Kawato and Gomi, 1992; Schweighofer et al., 1998; Tseng et al., 2007; Vinueza Veloz et al.,  
43 2015; Herzfeld et al. 2018). A fundamental question remains unanswered, however: how does  
44 the cerebellum learn to control the high dimensional and nonlinear motor systems that are  
45 typical of vertebrates for complicated movement patterns, while even the most advanced robots  
46 fail to perform similar movements (Adolph et al., 2012; Atkeson et al., 2018)? Moreover, how  
47 does the cerebellum achieve such learning despite being constrained by the relatively small  
48 numbers of training samples and the low-firing frequency from the inferior olive (IO) neurons,  
49 which give rise to the climbing fiber inputs? Here, we address these questions at three distinct  
50 levels: *computational*, *algorithmic*, and *implementational* (Marr, 1982), and provide  
51 computational and experimental clues for possible answers.

52

53 At the computational level, artificial learning systems that have many adjustable parameters  
54 require a proportionally large number of training samples to achieve adequate learning  
55 generalization (LeCun et al., 2015; Mnih et al., 2015; Silver et al. 2016). In contrast, if the  
56 number of training samples is lower than the number of parameters, severe overfitting to the  
57 noise in the data occurs. This creates a large generalization error, proportional to  $D/(2n)$ , where  
58  $n$  is the number of training samples and  $D$  is the number of degree-of-freedom (DOFs), which

59 is the number of adjustable parameters in the learning system (Watanabe, 2009; Yamazaki,  
60 2014). Therefore, an online learning system, such as the cerebellum, must keep the number of  
61 DOFs small in the early stages of learning when the training sample set is small, and then  
62 increase it gradually during the course of learning (Schaal et al., 2002; Garrigues and Ghaoui,  
63 2009).

64

65 At the algorithmic level, we earlier proposed that the number of cerebellar DOFs is modulated  
66 by the degree of synchrony between IO neurons (Kawato et al. 2011; Schweighofer et al. 2013;  
67 Tokuda et al. 2017). According to this proposal, early in learning, IO synchrony is high and  
68 groups of related neurons in the olivo-cerebellar system behave, in the limit, as a single-neuron  
69 chain, thus decreasing the number of DOFs. The resulting synchronous IO error signals would  
70 both significantly improve real-time motor control (Lang et al., 2016), and lead to massive  
71 changes in efficacies of the parallel-fiber-Purkinje-cell synapses, resulting in fast but crude  
72 learning. As learning of the motor act progresses, IO synchrony is decreased, potentially  
73 allowing the occurrence of chaotic resonance to enhance information transmission of the error  
74 signals (Schweighofer et al., 2004; Tokuda et al., 2010; Masuda and Aihara, 2002; Makarenko  
75 and Llinas, 2005; Nobukawa and Nishimura, 2016), which would overcome the constraint of low  
76 IO firing rates (Eccles et al., 1966; Llinás and Yarom, 1981). Specifically, chaotic resonance  
77 would increase the number of DOFs, and thereby allow more sophisticated learning, resulting in  
78 fine tuning of the motor act (Tokuda et al., 2010, 2013).

79

80 At the implementational level, we propose that the distinctive features of the IO neurons, the  
81 source of the climbing fibers, and the anatomy of the loop formed by the Purkinje cells, deep  
82 cerebellar nucleus neurons, and IO neurons, allow modulation of the number of DOFs during  
83 learning. The IO neurons form the strongest electrically coupled neuronal network in the adult

84 mammalian brain (De Zeeuw et al., 1995; Condorelli et al., 1998; Belluardo et al., 2000), with  
85 electrical synapses driving synchronization when the coupling is strong (Blenkinsop and Lang,  
86 2006 Marshall et al., 2007; Lang 2002). In addition, presynaptic GABAergic terminals control the  
87 efficacy of electrical coupling (Llinas et al., 1974; Sotelo et al., 1974; Best and Regehr, 2009;  
88 Onizuka et al., 2013; Lefler et al., 2014). These GABAergic afferents largely arise from the deep  
89 cerebellar nucleus (de Zeeuw et al., 1989; Nelson and Mugnaini, 1989; Fredette and Mugnaini,  
90 1991), which are part of the anatomical closed-loops formed between corresponding regions of  
91 the IO, cerebellar cortex, and deep cerebellar nuclei (Sugihara and Shinoda, 2004, 2007; Apps  
92 and Hawkes, 2009; Sugihara et al., 2009). Thus, the Purkinje cells, via this feedback circuit, can  
93 regulate the synchrony levels of their corresponding climbing fiber inputs, through the double-  
94 inhibition within the feedback circuit (Marshall and Lang, 2009).

95

96 Although we have shown in simulations that decreasing IO synchrony via modulation of  
97 electrical coupling enhances cerebellar learning (Schweighofer et al., 2004; Tokuda et al., 2010,  
98 2013), there are no experimental supports to the basic assumptions about how electrical  
99 coupling, synchrony, chaotic firing, and dimensionality of firing dynamics are linked. Here, we  
100 analyze the effect of coupling on the DOFs and on the induction of chaotic resonance by  
101 utilizing *in vivo* recordings of complex spikes under three coupling conditions. Specifically, we  
102 examine two predictions of our previous hypotheses that 1) increasing the synchrony level, via  
103 increased electrical coupling between inferior olive neurons, decreases the dimensionality of the  
104 IO firing dynamics and 2) intermediate coupling induces chaotic spiking and maximizes the  
105 dimensionality of inferior-olive firing dynamics.

106

107 **RESULTS**

108

109 **Estimation of the effective coupling between IO neurons *in vivo*.**

110 To examine the effect of electrical coupling on dimensionality reduction and chaotic dynamics,  
111 we first need to estimate the range of effective coupling strengths between IO neurons. Direct  
112 quantitative measurement of electrical coupling in the IO has been obtained in slice  
113 preparations (Devor and Yarom, 2002; Hoge et al., 2011; Lefler et al., 2014); however, it  
114 remains technically impossible to measure *in vivo*. We thus employed an indirect approach to  
115 estimate the coupling that involved comparing Purkinje cell complex spike activity recorded  
116 simultaneously from arrays of Purkinje cells (Blenkinsop and Lang, 2006; Lang, 2002; Lang et  
117 al., 1996) with simulated activity generated by a model of the IO using a Bayesian method,  
118 which we previously proposed and validated (Hoang et al., 2015)<sup>1</sup>. Here, we modified this  
119 method to further improve the robustness of the coupling estimations via Bayesian model-  
120 averaging.

121

122 Briefly, in the model, each IO neuron comprises a soma, a main dendrite, and four dendritic  
123 spine compartments, with these compartments having distinct ionic conductances. Most  
124 notably, the dendritic compartment has a high threshold calcium conductance and a calcium-  
125 activated potassium conductance, which are responsible for the after-depolarization and after-  
126 hyperpolarization sequence that follows each sodium spike and for the low firing rates of IO  
127 neurons (Schweighofer et al., 1999). Each neuron was coupled to its neighboring neurons via  
128 electric coupling conductances on the spine compartments, with one inhibitory input  
129 conductance per spine. Synaptic noise was added to better account for stochastic process in IO

---

<sup>1</sup> Purkinje cell complex spikes, as opposed to simple spikes that are due to the granule cell inputs, bear a one-to-one relationship to IO spikes. Thus, complex spikes can be used as a proxy for IO spikes (see Figure S1 for examples of complex spike recordings in these three conditions; see Methods for experimental procedures).

130 neurons - for review of IO anatomy and function, see (De Zeeuw et al., 1998) and of the model,  
131 see (Hoang et al., 2015; Onizuka et al., 2013).

132

133 From the model, it is possible to derive the theoretical "effective" electrical coupling  
134 conductance  $g_{eff}$  as a function of the axial conductance of the spines  $g_s$ , the electrical coupling  
135 conductance  $g_c$ , and the GABAergic synaptic conductance  $g_i$  (see Katori et al., 2010 and  
136 Experimental Procedures for details). Estimates for  $g_c$  and  $g_i$  were obtained by comparing the  
137 model spike activity to the complex spike data sets ( $g_s$  was held constant). However, initial  
138 simulations showed that the firing frequency of the synaptic noise inputs significantly affected  
139 the spiking behavior of the neurons in the model, thus the fit of the firing dynamics of the model  
140 to the data and the estimation results. To address this issue, we estimated  $g_c$  and  $g_i$  for different  
141 values of synaptic noise input frequencies via a model-averaging approach (Grueber et al.,  
142 2011). Specifically, we first constructed a number of models with different frequencies of  
143 synaptic noise inputs, as observed in cerebellar slice data (Najac and Raman, 2015). We then  
144 obtained an estimate of  $g_i$  and  $g_c$  via the Bayesian estimation method from each model (Hoang  
145 et al., 2015; see Experimental Procedures for details). The final estimates of  $g_i$  and  $g_c$  were  
146 obtained by averaging these individual model estimates, weighted in proportion to the  
147 goodness-of-fit of the models via Bayesian model-averaging (see Figure S2 and Experimental  
148 Procedures for details).

149

150 Following the procedures outlined above, estimates of  $g_i$  and  $g_c$  were obtained for three coupling  
151 conditions (low, control, high). The low coupling condition was generated by intra-IO injection of  
152 the gap junction blocker carbenoxolone (CBX), whereas the presumed high coupling condition  
153 was generated by intra-IO injection of the GABA blocker picrotoxin (PIX). Estimation results  
154 show that, as expected, the mean  $g_i$  and  $g_c$  were reduced approximately 20% and 22% under

155 the PIX and CBX from their own control (CON) values, respectively. When the two CON groups  
156 were combined, the estimated inhibitory conductance  $g_i$  in the CON condition (Figure 1A –  $1.11$   
157  $\pm 0.22$  mS/cm<sup>2</sup>,  $n = 90$  neurons) was significantly decreased in the PIX condition ( $0.84 \pm 0.26$   
158 mS/cm<sup>2</sup>,  $n = 46$  neurons, PIX-CON:  $p < 0.0001$ ) and was comparable to that in the CBX  
159 condition ( $1.02 \pm 0.13$  mS/cm<sup>2</sup>,  $n = 44$  neurons, CBX-CON:  $p = 0.03$ ). Similarly, the estimated  
160 gap-junctional conductance  $g_c$  in the CBX condition (Figure 1B –  $0.87 \pm 0.23$  mS/cm<sup>2</sup>, CBX-  
161 CON:  $p < 0.0001$ ) was smaller than in the CON condition ( $1.22 \pm 0.28$  mS/cm<sup>2</sup>), but there was  
162 no significant difference between the PIX ( $1.29 \pm 0.18$  mS/cm<sup>2</sup>, PIX-CON:  $p = 0.23$ ) and CON  
163 conditions.

164

165 As a result of these changes in  $g_i$  and  $g_c$ , the estimated effective coupling strength,  $g_{eff}$ ,  
166 calculated using Equation 1, differed across the three conditions (one-way ANOVA:  $p < 0.0001$ ).  
167 It was smallest for the CBX condition (Figure 1C –  $g_{eff} = 0.030 \pm 0.003$  mS/cm<sup>2</sup>, CBX-CON:  $p <$   
168  $0.0001$ ), intermediate for the CON condition ( $g_{eff} = 0.033 \pm 0.003$  mS/cm<sup>2</sup>) and largest for the  
169 PIX condition ( $g_{eff} = 0.037 \pm 0.002$  mS/cm<sup>2</sup>, PIX-CON:  $p < 0.0001$ ). The estimated  $g_i$  and  $g_c$   
170 parameters were then used in the neuronal network to generate simulated spike trains under all  
171 three conditions. In each case, the spike trains were comparable to those of the recorded  
172 complex-spike activity. In particular, the firing rates of neurons in the CBX condition were lower  
173 (see Figure 2A, firing rate of model  $0.58 \pm 0.55$  Hz; data  $0.59 \pm 0.55$  Hz) than in the CON  
174 condition (see Figure 2B, model  $1.34 \pm 0.77$  Hz; data  $1.36 \pm 0.8$  Hz), and firing rates in the PIX  
175 condition were higher (Figure 2C, model  $2.42 \pm 0.95$  Hz; data  $2.51 \pm 1.05$  Hz).

176

177 Next, to examine whether the estimates for effective coupling strengths in the three conditions  
178 were biologically realistic, we computed the coupling coefficients (CCs) for the model neurons.  
179 We first hyperpolarized all model neurons to  $-69$  mV by injection of  $I_{hyp} = -1$   $\mu A/cm^2$  to make

180 them responsive to the stimulus. We then injected a step current  $I_{cmd} = -1 \mu A/cm^2$  in the soma of  
181 the center neuron and computed the CCs as the average ratio of change in steady state  
182 membrane potentials of this “master” cell and its four neighboring cells in the network (Figure  
183 S3). As expected, CC was smaller in the CBX condition (Figure 1D –  $CC = 0.01 \pm 0.005$ , CBX-  
184 CON:  $p = 0.01$ ) and larger in the PIX condition ( $CC = 0.02 \pm 0.004$ , PIX-CON:  $p < 0.0001$ ) than  
185 in the CON condition ( $CC = 0.014 \pm 0.007$ ). Furthermore, CCs are highly compatible with  
186 previously reported *in vitro* values (Lefler et al., 2014). In particular, our results showed that CCs  
187 in the PIX and CON conditions were comparable to those of the control ( $CC = 0.021 \pm 0.02$ , cf.  
188 Table S1, Lefler et al., 2014) and light-activated ( $CC = 0.012 \pm 0.013$ ) conditions, respectively.  
189 In both cases, a larger (about double) CC value was found for the condition with less GABA  
190 activity (PIX in our experiments and control in Lefler et al, 2014).

191

## 192 **The dimensionality reduction by effective coupling**

193 To quantitatively investigate how synchrony and dimensionality change when effective coupling  
194 varies, we estimated the synchrony level and the dimensionality from the spike data of IO  
195 neurons in individual animals (see Experimental Procedures for more details). In strong  
196 agreement with previous studies (Blenkinsop and Lang 2006; Lang 1996), the synchrony level  
197 in 1-ms bins increased 2-3 fold in the PIX condition (synchrony =  $0.068 \pm 0.051$ , t-test PIX-CON:  
198  $p = 0.03$ , Fig 3A) and decreased about 70% in the CBX condition (synchrony =  $0.008 \pm 0.004$ , t-  
199 test CBX-CON:  $p = 0.04$ ) compared to the CON condition (synchrony =  $0.025 \pm 0.018$ ). In  
200 addition, when plotting the synchrony as a function of effective coupling averaged for each  
201 animal, we found a significant correlation, as expected (regression model in Wilkinson notation  
202 (Wilkinson and Rogers, 1973):  $synchrony \sim 1 + g_{eff}$ ,  $R^2 = 0.22$ , F-test:  $p = 0.04$ , Fig 3B) with a  
203 positive coefficient ( $mean \pm sem$ ,  $5 \pm 2$ ).

204

205 Next, we estimated the dimensionality ( $d$ ) of complex spike activity for selected neurons in each  
206 animal. Briefly, we extracted the average firing rates of neurons in 50-second long periods and  
207 applied principal component analysis (PCA) to compute the covariance of these firing rate  
208 vectors.  $d$  has been approximated (by Equation 5) as the minimal number of components  
209 accounting for 90% of variability in the data (Abbott et al., 2011). However, with this method, the  
210 dimensionality  $d$  is under-estimated for data with small numbers of neurons, as in our data  
211 (Mazzucato et al., 2015). We thus developed a new extrapolation method to correct the  
212 dimensionality estimation (see Experimental Procedures for more details). Results showed that  
213  $d$  was significantly reduced by PIX ( $d = 1.8 \pm 0.8$ , t-test PIX-CON:  $p = 0.02$ , Fig 3C) compared to  
214 the CON condition ( $d = 3.7 \pm 1.2$ ), but there was no significant difference between the CBX ( $d =$   
215  $3.5 \pm 1.6$ , t-test CBX-CON:  $p = 0.7$ ) and the CON conditions, probably because of a relatively  
216 small effect CBX had on complex spike firing rate. In addition, a regression analysis showed a  
217 strong correlation between  $g_{eff}$  and  $d$  (regression model:  $d \sim 1 + g_{eff}$ ,  $R^2 = 0.28$ , F-test:  $p = 0.025$ ,  
218 Fig 3D). Here the regression coefficient was negative ( $mean \pm sem$ ,  $-280 \pm 113$ ) and opposite to  
219 that of synchrony vs.  $g_{eff}$  ( $mean \pm sem$ ,  $5 \pm 2$ , see above), supporting our hypothesis that  
220 synchronization is a feasible mechanism for dimensionality reduction in IO neurons and that  
221 effective coupling is the control parameter for the IO to optimize the dimensionality of the olivo-  
222 cerebellar system.

223

#### 224 **Inverted U-shaped relationship between complexity entropy and effective coupling**

225 Finally, we addressed the question of whether physiological and intermediate coupling strengths  
226 maximize the chaotic level of IO activity. The Lyapunov exponents quantify the sensitivity of a  
227 dynamical system to initial conditions (Farmer and Sidorowich, 1987; Sano and Sawada, 1985),  
228 and are thus often used as indicators of chaos. However, methods to compute Lyapunov

229 exponents from time series data (Kantz, 1994; Rosenstein et al., 1993) are not applicable to our  
230 spike data sets, because the computation requires access to continuous variables. We therefore  
231 computed the complexity entropy (Letellier 2006; Hirata and Aihara, 2009, see Experimental  
232 Procedures for details), which has been shown to approximate the first Lyapunov exponents in  
233 simulations of the IO neurons (see Figure S6).

234

235 For both the simulated IO spike and the experimental complex spike data sets, we investigated  
236 whether the relationship between complexity entropy and effective coupling formed an inverted  
237 U-shape, as previously shown in simulations (Schweighofer et al., 2004; Tokuda et al., 2010).  
238 For each of the experimental IO neurons, we computed the complexity entropy from the  
239 simulated spike data that was generated with the estimated coupling values that best fit the data  
240 in terms of the PCA error (difference between experimental and simulated spike data in the PCA  
241 space, see Figure S2A). For the IO model, the second order model  $entropy \sim 1 + g_{eff} + g_{eff}^2$ ,  
242 where *entropy* is the complexity entropy, had a negative coefficient of the second order term  
243 (*mean ± sem*,  $-246 \pm 46$ ), and better fit the simulated spikes in the three conditions than the  
244 first-order linear model ( $entropy \sim 1 + g_{eff}$ ) (Figure 4A, the Log likelihood ratio (LLR):  $p <$   
245  $0.0001$ ). For the IO data (Figure 4B), a mixed effect regression model analysis, with *Animal* as a  
246 random intercept accounting for repeated measures within the same animal, showed that the  
247 second order model  $entropy \sim 1 + g_{eff} + g_{eff}^2 + (1 | Animal)$ , where  $(1 | Animal)$  is the random  
248 intercept, had a negative fixed-effect coefficient of the second order term (*mean ± sem*,  $-217 \pm$   
249  $63$ ), and provided a better fit than the linear model ( $entropy \sim 1 + g_{eff} + (1 | Animal)$ , LLR:  $p =$   
250  $0.0007$ ). Thus, for both the IO model (Figure 4A) and the experimental data (Figure 4B), an  
251 inverted U-curve that peaks at around  $g_{eff} = 0.032$  was found, indicating that similar intermediate  
252 coupling strengths induce chaotic behavior in both the model and the experimental data. It  
253 should be noted that the relatively small changes in the complexity entropy that we observed in

254 the model and data induce significant changes in firing dynamics, from synchronous and  
255 rhythmic firings ( $entropy = 0.24$ ,  $\lambda_1 = 20$  bits/second, cf. Figure S6B) to chaotic firings ( $entropy =$   
256  $0.29$ ,  $\lambda_1 = 50$  bits/second).

257

258

## 259 **DISCUSSION**

260

261 We developed a novel technique that combines computational modeling, Bayesian inference  
262 and model-averaging to estimate the effective coupling from rat *in vivo* complex spike data. The  
263 estimated effective coupling strengths of the three data conditions were consistent with the  
264 physiological effects of the drugs, i.e., increased by PIX and decreased by CBX. Notably, the  
265 coupling coefficients estimated from our simulations were highly compatible with *in vitro* values  
266 observed in (Lefler et al., 2014). In both studies, an approximate doubling of the CC was found  
267 for the conditions where GABA activity was lower or blocked. Such compatibility validates the  
268 estimation of coupling strengths in our study.

269

270 Our analysis of complex-spike data shows that increased electrical coupling between IO  
271 neurons decreases the dimensionality of the IO firing dynamics. Dimensionality reduction has  
272 long been considered one of the core computations in the brain (Pillow et al., 2008;  
273 Cunningham and Yu, 2014; Churchland et al. 2010; Rigotti et al., 2013; Mazzucato et al., 2016).  
274 Our study provides direct evidence that electrical coupling among neurons can control the  
275 dimensionality of the population activity by modulating the synchrony of the neural code. No  
276 significant difference of dimensionality between CBX and CON conditions was found, probably  
277 because of the incomplete effect of CBX in de-synchronizing complex spike activity (Blenkinsop

278 and Lang, 2006). Quantitatively, the approximately two-fold reduction in dimensionality from the  
279 PIX to the CON condition was highly comparable to that of stimuli-evoked activity of cortical  
280 neurons under different stimulus conditions and in varied tasks (Mazzucato et al., 2016;  
281 Churchland et al. 2010). We note, however, that additional mechanisms could work in parallel to  
282 effectively control the DOFs, such as pruning of irrelevant inputs (Cortese et al. 2018). In the  
283 olivo-cerebellar system, in particular, climbing fiber-Purkinje-cell synapses are gradually  
284 eliminated based on IO activity during development (Schweighofer, 1998; Good et al., 2017).  
285 Further experimental and computational analyses are required to elucidate the interplay  
286 between possible mechanisms in controlling the DOFs of the olivo-cerebellar system.

287

288 Our results also show that intermediate ranges of electrical coupling induce chaotic dynamics.  
289 In contrast, weak or strong coupling decrease the complexity entropy. The finding of an inverted  
290 U curve of complexity entropy as a function of effective coupling in both the model and  
291 experimental data are consistent with the “chaotic resonance” hypothesis, according to which  
292 chaotic firing increases information transmission despite the low firing rates of IO neurons  
293 (Schweighofer et al., 2004). We have previously proposed, and shown in simulations, that such  
294 chaotic firing may be useful to enhance cerebellar learning by increasing the error transmission  
295 capability of the olivocerebellar system (Tokuda et al., 2010). In agreement with this view, a  
296 previous study showed that the entropy of neural activity and mutual information between  
297 stimulus and response are maximized in balanced excitatory/inhibitory cortical networks (Shew  
298 et al., 2011).

299

300 Experimental data supporting the importance of electrical coupling for cerebellar learning comes  
301 from mice mutants which, due to lacking of electrotonic coupling between IO cells, exhibit  
302 deficits in learning-dependent motor tasks (Van Der Giessen et al., 2008). Similarly, humans

303 with reduced IO coupling show motor learning impairments (Van Essen et al., 2010). Because  
304 the inhibitory neurons controlling the strength of coupling between IO cells are largely located in  
305 the deep cerebellar nuclei (Fredette and Mugnaini, 1991; De Zeeuw et al., 1989), the major  
306 output station of the cerebellum, the strength of effective coupling, and thus the level of chaotic  
307 behavior, presumably depends on the modulation of the deep cerebellar nuclear neurons via  
308 plastic processes in the cerebellar cortex and nuclei (Best and Regehr, 2009; Chaumont et al.,  
309 2013; DeGrujil et al., 2014; Lefler et al., 2014; Turecek et al., 2014; Witter et al., 2013). Indeed,  
310 changes in simple spike levels produce significant changes in complex spike synchrony  
311 (Marshall and Lang, 2009). Thus, the IO coupling strength during cerebellar learning could be  
312 adaptively modulated, with the Purkinje cell-cerebellar nuclei-inferior olive triangle acting to  
313 decrease coupling along the progress of cerebellar learning (Kawato et al., 2011; Schweighofer  
314 et al., 2013; Tokuda et al., 2013).

315

316 According to this view, in the early phase of learning, the motor commands are strongly  
317 disturbed and far from the desired ones. The Purkinje cells, which are then strongly modulated  
318 by large sensory inputs and error signals, suppress the inhibitory effect of the neurons in  
319 cerebellar nuclei on the IO. Thus, the IO neurons are initially strongly coupled and the  
320 dimensionality is low. Because of this low dimensionality, the IO network would respond only to  
321 low-dimensional components of the error signals, which may convey only the gross features of  
322 the motor commands. However, the strong coupling allows a widespread synchrony among IO  
323 neurons and potentially leads to massive changes in the parallel-fiber-Purkinje-cell synaptic  
324 weights, resulting in fast but coarse learning. In addition to its effect on learning, such highly  
325 synchronized IO activity may have a downstream effect via a large network of synchronized  
326 Purkinje cells (Blenkinsop and Lang, 2011; Lang and Blenkinsop 2011; Tang et al., 2016), that  
327 could trigger an emergency or protective feedback motor commands in response to this error. In

328 contrast, in the late phase of learning, as the motor error becomes smaller, the Purkinje cell  
329 activity may become weaker, allowing increased activity of cerebellar nuclear neurons. This  
330 would result, in turn, in reduced IO coupling and higher dimensionality. At this stage, the  
331 moderate coupling strengths could induce chaotic IO spike activity that would transmit high  
332 dimensional error signals, resulting in more sophisticated learning (Shaikh et al., 2017). A  
333 further possibility is that these high dimensional signals would also be used for the fine grain  
334 motor control commands that are needed for precise motor coordination (Hoogland et al, 2015).

335 **ACKNOWLEDGEMENTS**

336 HH, KT and MK were supported by a contract with the National Institute of Information and  
337 Communications Technology entitled 'Development of network dynamics modeling methods for  
338 human brain data simulation systems'. HH and KT are partially supported by ImPACT Program  
339 of Council of Science, Technology and Innovation (Cabinet Office, Government of Japan). HH,  
340 ITT and KT are partially supported by MEXT Kakenhi (No. 17H06313). HH is partially supported  
341 by JST ERATO (JPMJER1801, "Brain-AI hybrid"). KA is partially supported by MEXT Kakenhi  
342 (No. 15H05707) and AMED under Grant Number JP18dm0307009. EJM acknowledges grants  
343 NIH NS095089 and NS37028. NS acknowledges grants NSF BCS-1031899 and  
344 1R56NS100528-01.

345 **REFERENCES**

346

347 Abbott, L. F., Rajan, K., and Sompolinsky, H. (2011). "Interactions between intrinsic and  
348 stimulus-evoked activity in recurrent neural networks," in *The Dynamic Brain: An Exploration of*  
349 *Neuronal Variability and its Functional Significance*, eds D. L. Glanzman and M. Ding (New  
350 York, NY: Oxford University Press), 65–82.

351 Adolph KE, Cole WG, Komati M, Garciaguirre JS, Badaly D, Lingeman JM, Chan GLY, Sotsky  
352 RB (2012). How do you learn to walk? Thousands of steps and dozens of falls per day. *Psychol.*  
353 *Sci.*, 23:1387–94.

354 Apps, R. and Hawkes, R. (2009). Cerebellar cortical organization: a one-map hypothesis. *Nat*  
355 *Rev Neurosci.* 10(9):670-81.

356 Atkeson CG, Benzun PWB, Banerjee N, Berenson D, Bove CP, Cui X, DeDonato M, Du R,  
357 Feng S, Franklin P, et al. (2018). What Happened at the DARPA Robotics Challenge Finals. In  
358 *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue*. Edited by Spenko  
359 M, Buerger S, Iagnemma K. Springer International Publishing: 667–684.

360 Bastian, A.J. (2006). Learning to predict the future: the cerebellum adapts feedforward  
361 movement control. *Curr. Opin. Neurobiol.* 16, 645–649.

362 Bazzigaluppi, P., Ruigrok, T., Saisan, P., de Zeeuw, C.I., and de Jeu, M. (2012). Properties of  
363 the Nucleo-Olivary Pathway: An In Vivo Whole-Cell Patch Clamp Study. *PLoS One* 7.

364 Belluardo, N., Mudò, G., Trovato-Salinaro, A., Le Gurun, S., Charollais, A., Serre-Beinier, V.,  
365 Amato, G., Haefliger, J.A., Meda, P. and Condorelli D.F. (2000). Expression of connexin36 in  
366 the adult and developing rat brain. *Brain Res.* 865(1):121-38.

367 Best, A.R., and Regehr, W.G. (2009). Inhibitory Regulation of Electrically Coupled Neurons in

- 368 the Inferior Olive Is Mediated by Asynchronous Release of GABA. *Neuron* 62, 555–565.
- 369 Blenkinsop, T.A., and Lang, E.J. (2006). Block of Inferior Olive Gap Junctional Coupling  
370 Decreases Purkinje Cell Complex Spike Synchrony and Rhythmicity. *J. Neurosci.* 26, 1739–  
371 1748.
- 372 Blenkinsop, T. A., and Lang, E. J. (2011). Synaptic action of the olivocerebellar system on  
373 cerebellar nuclear spike activity. *J Neurosci.* 31(41), 14708-14720.
- 374 Chaumont, J., Guyon, N., Valera, A.M., Dugue, G.P., Popa, D., Marcaggi, P., Gautheron, V.,  
375 Reibel-Foisset, S., Dieudonne, S., Stephan, A., et al. (2013). Clusters of cerebellar Purkinje  
376 cells control their afferent climbing fiber discharge. *Proc. Natl. Acad. Sci.* 110, 16223–16228.
- 377 Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT,  
378 Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM,  
379 Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, Finn IM, Ferster D, Ryu SI,  
380 Santhanam G, Sahani M, Shenoy KV. (2010). Stimulus onset quenches neural variability: a  
381 widespread cortical phenomenon. *Nat Neurosci.* 13(3):369-78
- 382 Condorelli, D.F., Parenti, R., Spinella, F., Trovato Salinaro, A., Belluardo, N., Cardile, V. and  
383 Cicirata F. (1998). Cloning of a new gap junction gene (Cx36) highly expressed in mammalian  
384 brain neurons. *Eur J Neurosci.* 10(3):1202-8.
- 385 Cortese A., De Martino B. and Kawato M. (2018). The neural and cognitive architecture for  
386 learning from a small sample, [arXiv.org,1810.02476](https://arxiv.org/abs/1810.02476).
- 387 Cunningham J.P. & Yu B.M. (2014). Dimensionality reduction for large-scale neural recordings.  
388 *Nature Neurosci.* 17:1500–1509
- 389 D'Angelo, E., Mapelli, L., Casellato, C., Garrido, J.A., Luque, N., Monaco, J., Prestori, F.,  
390 Pedrocchi, A., and Ros, E. (2016). Distributed Circuit Plasticity: New Clues for the Cerebellar  
391 Mechanisms of Learning. *Cerebellum* 15, 139–151.

- 392 DeGruijl, J.R., Sokół, P.A., Negrello, M., and DeZeeuw, C.I. (2014). Modulation of electrotonic  
393 coupling in the inferior olive by inhibitory and excitatory inputs: Integration in the glomerulus.  
394 *Neuron* *81*, 1215–1217.
- 395 Devor, A., and Yarom, Y. (2002). Electrotonic coupling in the inferior olivary nucleus revealed by  
396 simultaneous double patch recordings. *J. Neurophysiol.* *87*, 3048–3058.
- 397 Eccles, J.C., Llinás, R., and Sasaki, K. (1966). The excitatory synaptic action of climbing fibres  
398 on the purkinje cells of the cerebellum. *J. Physiol.* *182*, 268–296.
- 399 Eckmann, J.-P., Kamphorst, S.O., and Ruelle, D. (1987). Recurrence Plots of Dynamical  
400 Systems. *Europhys. Lett.* *4*, 973–977.
- 401 Van Essen, T.A., Van der Giessen, R.S., Koekkoek, S.K.E., VanderWerf, F., De Zeeuw, C.I.,  
402 van Genderen, P.J.J., Overbosch, D., and de Jeu, M.T.G. (2010). Anti-malaria drug mefloquine  
403 induces motor learning deficits in humans. *Front. Neurosci.* *4*.
- 404 Farmer, J.D., and Sidorowich, J.J. (1987). Predicting chaotic time series. *Phys. Rev. Lett.* *59*,  
405 845–848.
- 406 Fredette, B.J., and Mugnaini, E. (1991). The GABAergic cerebello-olivary projection in the rat.  
407 *Anat. Embryol. (Berl).* *184*, 225–243.
- 408 Van Der Giessen, R.S., Koekkoek, S.K., van Dorp, S., De Gruijl, J.R., Cupido, A., Khosrovani,  
409 S., Dortland, B., Wellershaus, K., Degen, J., Deuchars, J., et al. (2008). Role of Olivary  
410 Electrical Coupling in Cerebellar Motor Learning. *Neuron* *58*, 599–612.
- 411 Garrigues, P. J. and Ghaoui L. E. (2009). An homotopy algorithm for the Lasso with online  
412 observations. *Advances in neural information processing systems*, NIPS 2009.
- 413 Good J-M., Mahoney M., Miyazaki T., Tanaka K. F., Sakimura K., Watanabe M., Kitamura K.  
414 and Kano M. (2017). Maturation of Cerebellar Purkinje Cell Population Activity during Postnatal

- 415 Refinement of Climbing Fiber Network. *Cell Reports*, 21(8):2066 - 2073.
- 416 Grueber, C.E., Nakagawa, S., Laws, R.J., and Jamieson, I.G. (2011). Multimodel inference in  
417 ecology and evolution: Challenges and solutions. *J. Evol. Biol.* 24, 699–711.
- 418 Hansel, C., Linden, D.J., and D’Angelo, E. (2001). Beyond parallel fiber LTD: the diversity of  
419 synaptic and non-synaptic plasticity in the cerebellum. *Nat. Neurosci.* 4, 467–475.
- 420 Herzfeld, D. J., Kojima, Y., Soetedjo, R. and Shadmehr, R. (2018). Encoding of error and  
421 learning to correct that error by the Purkinje cells of the cerebellum. *Nat. Neurosci.*, 21: 736–  
422 743.
- 423 Hirata, Y., and Aihara, K. (2009). Representing spike trains using constant sampling intervals. *J.*  
424 *Neurosci. Methods* 183, 277–286.
- 425 Hirata, Y., Horai, S., and Aihara, K. (2008). Reproduction of distance matrices and original time  
426 series from recurrence plots and their applications. *Eur. Phys. J. Spec. Top.* 164, 13–22.
- 427 Hoang, H., Yamashita, O., Tokuda, I.T., Sato, M.-A., Kawato, M., and Toyama, K. (2015).  
428 Segmental Bayesian estimation of gap-junctional and inhibitory conductance of inferior olive  
429 neurons from spike trains with complicated dynamics. *Front. Comput. Neurosci.* 9, 1–14.
- 430 Hoge, G.J., Davidson, K.G., Yasumura, T., Castillo, P.E., Rash, J.E., and Pereda, A.E. (2011).  
431 The extent and strength of electrical coupling between inferior olivary neurons is  
432 heterogeneous. *J. Neurophysiol.* 105, 1089–1101.
- 433 Hoogland, T.M., De Gruijl, J.R., Witter, L., Canto, C.B., De Zeeuw, C.I. (2015). Role of  
434 synchronous activation of cerebellar Purkinje cell ensembles in multi-joint movement control.  
435 *Curr Biol.* 25:1157–65.
- 436 Ito, M. (2001). Cerebellar long-term depression: characterization, signal transduction, and  
437 functional roles. *Physiol. Rev.* 81, 1143–1195.

- 438 Kantz, H. (1994). A robust method to estimate the maximal Lyapunov exponent of a time series.  
439 *Phys. Lett. A* *185*, 77–87.
- 440 Kaplan, J.L., and Yorke, J.A. (1970). Chaotic behavior of multidimensional difference equations.  
441 In H. O. Walter & H. Peitgen (Eds.). In *Lecture Notes in Mathematics: Vol. 730. Functional*  
442 *Differential Equations and Approximations of Fixed Points*, pp. 204–207.
- 443 Katori, Y., Lang, E.J., Onizuka, M., Kawato, M., and Aihara, K. (2010). Quantitative Modeling of  
444 Spatio-Temporal Dynamics of Inferior Olive Neurons With a Simple Conductance-Based Model.  
445 *Int. J. Bifurcat. Chaos* *20*, 583–603.
- 446 Kawato, M., and Gomi, H. (1992). The cerebellum and VOR/OKR learning models. *Trends*  
447 *Neurosci.* *15*, 445–453.
- 448 Kawato, M., Kuroda, S., and Schweighofer, N. (2011). Cerebellar supervised learning revisited:  
449 Biophysical modeling and degrees-of-freedom control. *Curr. Opin. Neurobiol.* *21*, 791–800.
- 450 Keating, J.G., and Thach, W.T. (1995). Nonclock behavior of inferior olive neurons: interspike  
451 interval of Purkinje cell complex spike discharge in the awake behaving monkey is random. *J.*  
452 *Neurophysiol.* *73*, 1329–1340.
- 453 Kitazawa, S., Kimura, T., and Yin, P.B. (1998). Cerebellar complex spikes encode both  
454 destinations and errors in arm movements. *Nature* *392*, 494–497.
- 455 Kobayashi, Y., Kawano, K., Takemura, A., Inoue, Y., Kitama, T., Gomi, H., and Kawato, M.  
456 (1998). Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during  
457 ocular following responses in monkeys II. Complex spikes. *J. Neurophysiol.* *80*, 832–848.
- 458 Kuroda, S., Schweighofer, N., and Kawato, M. (2001). Exploration of signal transduction  
459 pathways in cerebellar long-term depression by kinetic simulation. *J. Neurosci.* *21*, 5693–5702.
- 460 Lang, E.J. (2002). GABAergic and glutamatergic modulation of spontaneous and motor-cortex-

- 461 evoked complex spike activity. *J. Neurophysiol.* *87*, 1993–2008.
- 462 Lang E.J., Apps R., Bengtsson F., Cerminara N.L., De Zeeuw C.I., Ebner T.J., Heck D.H.,  
463 Jaeger D., Jörntell H., Kawato M., Otis T.S., Ozyildirim O., Popa L.S., Reeves A.M.B.,  
464 Schweighofer N., Sugihara I., Xiao J. (2016). The roles of the olivocerebellar pathway in motor  
465 learning and motor control, *The Cerebellum*, *16*(1): 230-252.
- 466 Lang, E. J., and Blenkinsop, T. A. (2011). Control of cerebellar nuclear cells: a direct role for  
467 complex spikes? *Cerebellum* *10*: 694-701.
- 468 Lang, E.J., Sugihara, I., and Llinás, R. (1996). GABAergic modulation of complex spike activity  
469 by the cerebellar nucleoolivary pathway in rat. *J. Neurophysiol.* *76*, 255–275.
- 470 LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature*, *521*:436–444.
- 471 Lee, K. H., Mathews, P. J., Reeves, A. M., Choe, K. Y., Jami, S. A., Serrano, R. E. and Otis, T.  
472 S. (2015). Circuit mechanisms underlying motor memory formation in the cerebellum. *Neuron*  
473 *86*(2): 529-40
- 474 Lefler, Y., Yarom, Y., and Uusisaari, M. (2014). Cerebellar inhibitory input to the inferior olive  
475 decreases electrical coupling and blocks subthreshold oscillations. *Neuron* *81*, 1389–1400.
- 476 Letellier, C. (2006). Estimating the Shannon entropy: Recurrence plots versus symbolic  
477 dynamics. *Phys. Rev. Lett.* *96*, 254102.
- 478 Llinás R. The noncontinuous nature of movement execution. In: Humphrey DR, Freund H-J,  
479 editors. *Motor control: concepts and issues*. New York: Wiley; 1991. p. 223–42.
- 480 Llinás, R., Baker R., and Sotelo C. (1974). Electrotonic coupling between neurons in cat inferior  
481 olive. *J. Neurophysiol.* *37*(3), 560-571.
- 482 Llinás, R., and Yarom, Y. (1981). Electrophysiology of mammalian inferior olivary neurones in  
483 vitro. Different types of voltage-dependent ionic conductances. *J. Physiol.* *315*, 549–567.

- 484 Long M.A., Deans M.R., Paul D.L., Connors B.W. (2002) Rhythmicity without synchrony in the  
485 electrically uncoupled inferior olive. *J Neurosci.* 22(24):10898-905.
- 486 Makarenko, V., and Llinas, R. (1998). Experimentally determined chaotic phase synchronization  
487 in a neuronal system. *Proc. Natl. Acad. Sci. United States Am. Sci.* 95, 15747–15752.
- 488 Marr, D. (1969). A theory of cerebellar cortex. *J Physiol.* 202(2): 437–470.
- 489 Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and*  
490 *Processing of Visual Information.* Henry Holt, New York, NY, USA.
- 491 Marshall S.P., van der Giessen R.S., de Zeeuw C.I. and Lang E.J. (2007). Altered  
492 olivocerebellar activity patterns in the connexin36 knockout mouse. *Cerebellum.* 6(4):287-99.
- 493 Marshall, S. P. and Lang E. J. (2009). Local changes in the excitability of the cerebellar cortex  
494 produce spatially restricted changes in complex spike synchrony. *J Neurosci* 29(45), 14352-  
495 14362.
- 496 Marwan, N., Carmen Romano, M., Thiel, M., and Kurths, J. (2007). Recurrence plots for the  
497 analysis of complex systems. *Phys. Rep.* 438, 237–329.
- 498 Masuda, N., and Aihara, K. (2002). Spatiotemporal spike encoding of a continuous external  
499 signal. *Neural Comput.* 14, 1599–1628.
- 500 Mazzucato L, Fontanini A and La Camera G (2016) Stimuli Reduce the Dimensionality of  
501 Cortical Activity. *Front. Syst. Neurosci.* 10:11. doi: 10.3389/fnsys.2016.00011
- 502 Mnih V, Kavukcuoglu K, Silver D, Rusu A a, Veness J, Bellemare MG, Graves A, Riedmiller M,  
503 Fidjeland AK, Ostrovski G, et al. (2015). Human-level control through deep reinforcement  
504 learning. *Nature*, 518:529–533.
- 505 Najac, M., and Raman I. M. (2015). Integration of Purkinje Cell Inhibition by Cerebellar Nucleo-  
506 Olivary Neurons. *J. Neurosci.* 35(2), 544–549.

- 507 Nelson, B. J. and Mugnaini, E. (1989). Origins of GABAergic inputs to the inferior olive. In  
508 Strata, P. (ed.), *The Olivocerebellar System in Motor Control*. Springer-Verlag. Berlin. Exp.  
509 Brain Res., Suppl. 17, 86-107.
- 510 Nobukawa, S., and Nishimura, H. (2016). Chaotic resonance in coupled inferior olive neurons  
511 with the Llinás approach neuron model. *Neural Comput.* 28, 2505–2532.
- 512 Onizuka, M., Hoang, H., Kawato, M., Tokuda, I.T., Schweighofer, N., Katori, Y., Aihara, K.,  
513 Lang, E.J., and Toyama, K. (2013). Solution to the inverse problem of estimating gap-junctional  
514 and inhibitory conductance in inferior olive neurons from spike trains by network model  
515 simulation. *Neural Networks* 47, 51–63.
- 516 Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E. J. and Simoncelli, E.  
517 P. (2008). Spatio-temporal correlations and visual signaling in a complete neuronal population.  
518 *Nature* 454: 995-999.
- 519 Rigotti M., Barak O., Warden M.R., Wang X.J., Daw N.D, Miller E.K. & Fusi S. (2013) The  
520 importance of mixed selectivity in complex cognitive tasks. *Nature*. 497:585–590.
- 521 Rosenstein, M.T., Collins, J.J., and De Luca, C.J. (1993). A practical method for calculating  
522 largest Lyapunov exponents from small data sets. *Phys. D* 65, 117–134.
- 523 Sano, M., and Sawada, Y. (1985). Measurement of the lyapunov spectrum from a chaotic time  
524 series. *Phys. Rev. Lett.* 55, 1082–1085.
- 525 Schaal, S., Atkeson, C.G. and Vijayakumar, S. (2002). Scalable Techniques from  
526 Nonparametric Statistics for Real Time Robot Learning. *Applied Intelligence*. 17: 49-60.
- 527 Schweighofer N. (1998). A model of activity-dependent formation of cerebellar microzones.  
528 *Biological Cybernetics*, 79 (2): 97-107.
- 529 Schweighofer, N., Spoelstra, J., Arbib, M.A., and Kawato, M. (1998). Role of the cerebellum in

- 530 reaching movements in humans. II. A neural model of the intermediate cerebellum. *Eur. J.*  
531 *Neurosci.* *10*, 95–105.
- 532 Schweighofer, N., Doya, K., and Kawato, M. (1999). Electrophysiological properties of inferior  
533 olive neurons: A compartmental model. *J. Neurophysiol.* *82*, 804–817.
- 534 Schweighofer, N., Doya, K., Fukai, H., Chiron, J.V., Furukawa, T., and Kawato, M. (2004).  
535 Chaos may enhance information transmission in the inferior olive. *Proc. Natl. Acad. Sci.* *101*,  
536 4655–4660.
- 537 Schweighofer, N., Lang, E.J., and Kawato, M. (2013). Role of the olivo-cerebellar complex in  
538 motor learning and control. *Front. Neural Circuits* *7*, 94.
- 539 Shaikh, A. G., Wong, A. L., Optican, L. M. and Zee, D. S. (2017) Impaired Motor Learning in a  
540 Disorder of the Inferior Olive: Is the Cerebellum Confused? *Cerebellum* *16*: 158-167.
- 541 Shew W.L., Yang H., Yu S., Roy R. and Plenz D. (2011) Information Capacity and Transmission  
542 Are Maximized in Balanced Cortical Networks with Neuronal Avalanches. *J. Neurosci.* *31*(1) 55-  
543 63.
- 544 Shinomoto, S., Miura, K., and Koyama, S. (2005). A measure of local variation of inter-spike  
545 intervals. *BioSystems*, *79*(1-3): 67-72.
- 546 Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J,  
547 Antonoglou I, Panneershelvam V, Lanctot M, et al. (2016). Mastering the game of Go with deep  
548 neural networks and tree search. *Nature*, *529*:484–489.
- 549 Sotelo, C., Llinás R., and Baker R. (1974). Structural study of inferior olivary nucleus of the cat:  
550 morphological correlates of electrotonic coupling. *J. Neurophysiol.* *37*(3), 541-559.
- 551 Sugihara, I., Fujita, H., Na, J., Quy, P.N., Li, B.Y. and Ikeda, D. (2009). Projection of  
552 reconstructed single Purkinje cell axons in relation to the cortical and nuclear aldolase C

- 553 compartments of the rat cerebellum. *J Comp Neurol.* 512(2):282-304.
- 554 Sugihara, I. and Shinoda, Y. (2004). Molecular, topographic, and functional organization of the  
555 cerebellar cortex: a study with combined aldolase C and olivocerebellar labeling. *J Neurosci.*  
556 24(40): 8771-85.
- 557 Sugihara, I. and Shinoda, Y. (2007). Molecular, Topographic, and Functional Organization of the  
558 Cerebellar Nuclei: Analysis by Three-Dimensional Mapping of the Olivonuclear Projection and  
559 Aldolase C Labeling. *J Neurosci.* 27(36): 9696-9710.
- 560 Tang, T., Suh C. Y., Blenkinsop T. A., and Lang E. J. (2016). Synchrony is Key: Complex Spike  
561 Inhibition of the Deep Cerebellar Nuclei. *Cerebellum* 15(1), 10-13.
- 562 Tokuda, I.T., Han, C.E., Aihara, K., Kawato, M., and Schweighofer, N. (2010). The role of  
563 chaotic resonance in cerebellar learning. *Neural Networks* 23, 836–842.
- 564 Tokuda, I.T., Hoang, H., Schweighofer, N., and Kawato, M. (2013). Adaptive coupling of inferior  
565 olive neurons in cerebellar learning. *Neural Networks* 47, 42–50.
- 566 Tokuda, I.T., Hoang, H., and Kawato, M. (2017). New insights into olivo-cerebellar circuits for  
567 learning from a small training sample, *Curr. Opin. Neurobiol.* 46:58-67.
- 568 Tseng, Y.-W., Diedrichsen, J., Krakauer, J.W., Shadmehr, R., and Bastian, A.J. (2007). Sensory  
569 prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.* 98, 54–  
570 62.
- 571 Turecek, J., Yuen, G.S., Han, V.Z., Zeng, X.H., Bayer, K.U., and Welsh, J.P. (2014). NMDA  
572 receptor activation strengthens weak electrical coupling in mammalian brain. *Neuron* 81, 1375–  
573 1388.
- 574 Victor, J.D., and Purpura, K.P. (1997). Metric-space analysis of spike trains: theory, algorithms  
575 and application. *Netw. Comput. Neural Syst.* 8, 127–164.

- 576 Vinueza Veloz, M.F., Zhou, K., Bosman, L.W.J., Potters, J.W., Negrello, M., Seepers, R.M.,  
577 Strydis, C., Koekkoek, S.K.E., and De Zeeuw, C.I. (2015). Cerebellar control of gait and  
578 interlimb coordination. *Brain Struct. Funct.* *220*, 3513–3536.
- 579 Watanabe S (2009). *Algebraic Geometry and Statistical Learning Theory*. Cambridge University  
580 Press.
- 581 Welsh, J.P., Lang, E.J., Sugihara, I., Llinás, R. (1995). Dynamic organization of motor control  
582 within the olivocerebellar system. *Nature* *374*:453–7.
- 583 Wilkinson, G.N., and Rogers, C.E. (1973). Symbolic description of factorial models for analysis  
584 of variance. *Appl. Stat.* *22*, 392–399.
- 585 Witter, L., Canto, C.B., Hoogland, T.M., de Gruijl, J.R., and De Zeeuw, C.I. (2013). Strength and  
586 timing of motor responses mediated by rebound firing in the cerebellar nuclei after Purkinje cell  
587 activation. *Front. Neural Circuits* *7*: 133.
- 588 Wolf, A., Swift, J.B., Swinney, H.L., and Vastano, J.A. (1985). Determining Lyapunov exponents  
589 from a time series. *Phys. D Nonlinear Phenom.* *16*, 285–317.
- 590 Yamazaki K (2014). Asymptotic accuracy of distribution-based estimation of latent variables. *J*  
591 *Mach Learn Res.* *15*:3721-3742.
- 592 De Zeeuw, C.I., and Ten Brinke, M.M. (2015). Motor learning and the cerebellum. *Cold Spring*  
593 *Harb. Perspect. Biol.* *7*.
- 594 De Zeeuw, C.I., Holstege, J.C., Ruigrok, T.J.H., and Voogd, J. (1989). Ultrastructural study of  
595 the GABAergic, cerebellar, and mesodiencephalic innervation of the cat medial accessory olive:  
596 Anterograde tracing combined with immunocytochemistry. *J. Comp. Neurol.* *284*, 12–35.
- 597 De Zeeuw, C.I., Hertzberg, E.L., and Mugnaini, E. (1995). The dendritic lamellar body: a new  
598 neuronal organelle putatively associated with dendrodendritic gap junctions. *J. Neurosci.* *15*,

599 1587–1604.

600 De Zeeuw, C.I., Simpson, J.I., Hoogenraad, C.C., Galjart, N., Koekkoek, S.K.E., and Ruigrok,

601 T.J.H. (1998). Microcircuitry and function of the inferior olive. *Trends Neurosci.* 21, 391–400.

602

## 603 **EXPERIMENTAL PROCEDURES**

604

605 The recording experiments were performed in accordance with the National Institute of Health  
606 Guide for the Care and Use of Laboratory Animals. Experimental protocols were approved by  
607 the Institutional Animal Care and Use Committee of New York University School of Medicine.

608

### 609 **Experimental data**

610 The analyses were performed on a subset of data obtained in two prior series of experiments in  
611 ketamine/xylazine anesthetized female, Sprague-Dawley rats that involved either injection of  
612 picrotoxin (PIX) or carbenoxolone (CBX) to the IO to block GABA-A receptors or gap junctions,  
613 respectively (Blenkinsop and Lang, 2006; Lang, 2002; Lang et al., 1996). The specific  
614 experiments were chosen primarily on the basis of having typical complex spike activity in  
615 control and a large change in activity in response to the drug injection.

616

617 Details of the experimental procedures can be found the original reports. In brief, a rectangular  
618 array of glass microelectrodes was implanted into the apical surface of crus 2a. The arrays  
619 typically contained 3-4 mediolaterally running rows and up to 10 rostrocaudally running  
620 columns, with an interelectrode spacing of  $\sim 250 \mu\text{m}$ . Electrodes were implanted to a depth of  
621  $\sim 100 \mu\text{m}$  below the brain surface such that complex spikes from individual Purkinje cells were  
622 recorded. In each experiment, spontaneous complex spike activity was recorded during an initial  
623 control period. Following the control (CON) period, the IO was located by lowering a  
624 microelectrode through the brainstem under stereotaxic guidance until activity characteristic of  
625 IO neurons was observed. The microelectrode was then replaced by an injection pipette  
626 containing the drug solution that was lowered to the same location as the site where IO activity

627 was found. A slow injection of drug solution was then performed ( $\sim 1 \mu\text{l}$  over 5-10 min). The drug  
628 conditions analyzed were recorded after completion of the injection and a clear change in  
629 activity was observed. The multielectrode arrays recorded from 10 - 30 Purkinje cells in each of  
630 the CBX experiments (n = 6 animals), and from 16- 42 Purkinje cells in the PIX experiments (n =  
631 3 animals).

632

633 The effect of CBX and PIX on complex spike activity often varied among cells within an  
634 experiment. This was likely due to the Purkinje cells in different parts of the array receiving  
635 climbing fibers from different regions of the IO, that the drugs were injected at a single point  
636 within the IO, and that drug concentration (and therefore potentially efficacy) will fall with  
637 distance from the injection site. Indeed, the IO is an extended structure (particularly in the  
638 rostrocaudal axis where it is  $\sim 2$  mm long). We therefore considered the effects of the drugs  
639 when selecting the neurons for analysis. That is, Purkinje cells that exhibited significant changes  
640 in complex spike synchrony, measured as the coincidence of spikes in 1 ms time bins, between  
641 the control and drug conditions were selected. For CBX, the criterion was a 50% decrease and  
642 for PIX it was a 200% increase. In total, we analyzed spike train data from 500-second long  
643 periods for the control and drug conditions for each neuron (neurons/condition: control, n = 90;  
644 PIX, n = 46; CBX, n= 44).

645

#### 646 **IO network model**

647 The IO neuron model is a conductance-based model (Schweighofer et al., 1999) extended via  
648 addition of glomerular compartments comprising electrically coupled spines (Onizuka et al.,  
649 2013). The network model consisted of an array of 3x3 IO neurons, each of which was mutually  
650 connected to its four neighboring neurons by a gap junction from one of its spines to one of its

651 neighbor's represented by the gap-junctional conductance  $g_c$ . We simulated spike data of the  
652 nine cells with step-wise changes of two model parameters, i.e., inhibitory synaptic conductance  
653  $g_i$ , and coupling conductance  $g_c$ . These two parameters were both varied in the range of 0–2.0  
654 mS/cm<sup>2</sup> with an increment of 0.05 mS/cm<sup>2</sup>. We generated a total of 41x41=1681 sets of 500-  
655 second long simulated spike trains. The simulated spike data for each variation of  $g_i$  and  $g_c$  was  
656 then compared with the actual spike data, and the parameters whose firing dynamics best fit to  
657 that of individual neurons in the control, PIX, and CBX conditions were selected as the  
658 estimated values. Because the effect of the axial conductance of the spines,  $g_s$ , is equivalent to  
659 that of the gap-junctional conductance,  $g_c$ , in determining the amount of current will flow across  
660 the gap junction,  $g_s$  does not need to be estimated from the data and thus was fixed at 0.1  
661 mS/cm<sup>2</sup> (Onizuka et al., 2013). To better account for excitability of the neurons *in vivo*, the  
662 inward sodium current conductance  $g_{Na}$  was set as 110 mS/cm<sup>2</sup>, which has been shown to  
663 induce robust chaos in the model (Schweighofer et al., 2004).

664

### 665 **The segmental Bayes inference for estimating the effective coupling from a single model**

666 Under simplified assumptions, the effective coupling,  $g_{eff}$ , between two IO neurons was  
667 calculated from the axial conductance of the spines  $g_s$ , inhibitory conductance  $g_i$  and gap-  
668 junctional conductance  $g_c$  as in (Katori et al., 2010):

$$669 \quad g_{eff} = \frac{g_s}{2g_c + g_i + g_s} g_c. \quad (1)$$

670 This equation implies that to estimate the effective coupling  $g_{eff}$ , we need to estimate both the  
671 coupling conductance  $g_c$  and the GABA conductance  $g_i$  reliably for each of the three datasets  
672 CBX, CON, and PIX. For that purpose, we previously developed a Bayesian method that  
673 contains two steps (see Supplemental Materials for outlines of the Bayesian method, and  
674 Hoang et al., 2015). In the first step, the parameters are estimated for each 50-second time-

675 segment of individual neurons, allowing the parameter values to vary in time. This compensates  
676 for inevitable mismatch in the firing patterns between the model and the data. In the second  
677 step, a single set of parameter values is estimated for the entire time-segments of individual  
678 neurons by a hierarchical Bayes framework. Below, we outline the segmental Bayes method (for  
679 a detailed description see, Hoang et al., 2015).

680

681 First, the firing dynamics of the spike data were characterized by a feature vector composed of  
682 a total of sixty-eight spatiotemporal features, e.g., firing rate, local variation (Shinomoto et al.,  
683 2005), cross-correlation, auto-correlation, and minimal distance (Hirata and Aihara, 2009).  
684 Principal component analysis (PCA) was then conducted to remove the redundancy of those  
685 features. The Bayesian inference aims to inversely estimate the conductance values from the  
686 top three-dimensional principal components, which accounted for 55% of the data variance. To  
687 compensate for the modeling errors, i.e. differences in the complexity of firing patterns between  
688 the model and actual neurons, we divided the spike data of each neuron into short time-  
689 segments under the assumption that segmental estimates of individual neurons fluctuated  
690 around a single neuronal estimate with a normal (Gaussian) distribution. The conductance  
691 values of individual neurons can be estimated by a hierarchical Bayesian framework. Here, the  
692 segment size, 50 seconds, was optimized so that the variance of firing frequency across  
693 segments was minimal (Onizuka et al., 2013). In addition, we also introduced two physiological  
694 constraints on the estimates: a common  $g_i$  for CBX and CON neurons in CBX experiments and  
695 a common  $g_c$  for PIX and CON neurons in PIX experiments. The rationale for these constraints  
696 is that CBX and PIX are supposed to only reduce the gap-junctional conductance  $g_c$  and  
697 inhibitory conductance  $g_i$ , respectively.

698

699 We have shown that the segmental Bayes algorithm minimizes the fitting between experimental

700 and simulated spike data (Hoang et al., 2015), and further confirmed, by simulations, that it  
701 indeed minimizes the estimation errors compared to other conventional methods – including the  
702 non-segmental Bayes inference, which finds the estimates once across the entire spike data,  
703 and the minimum-error algorithm, which directly finds the closest match in the feature space  
704 (data not shown).

705

### 706 **Model-averaging estimation of the effective coupling**

707 We found that the firing frequency of inhibitory synaptic noise inputs largely affect spiking  
708 behavior of the IO model and thus the estimation results. To reduce the uncertainty in estimates  
709 of  $g_c$  and  $g_i$ , we therefore adopted the segmental Bayes algorithm by a model-averaging  
710 approach as follows (for review, see Grueber et al., 2011). Due to the extremely expensive  
711 computation of the compartmental model (about a week for ten computer clusters to generate  
712 the spiking data of 1681 conductance values of a single model), we first simulated four models  
713 with the firing frequency of inhibitory synaptic inputs of 10, 20, 50 and 70 Hz, which are  
714 observed in slices of cerebellar nucleo-olivary neurons (Najac and Raman, 2015). Next, we  
715 conducted the segmental Bayes to estimate posterior probability of  $g_i$  and  $g_c$  for each model.

$$716 \quad P(g|y, m_i) \propto P(y|g, m_i)P(g), \quad (2)$$

717 where  $P(g | y, m_i)$  is the posterior probability of the conductance  $g = (g_i, g_c)$ ,  $y$  is the feature  
718 vectors extracted from the spike data, and  $m_i$  is the  $i$ th selected model ( $i = 1..4$ ). We then mixed  
719 the posterior probabilities with the weights proportional to the model evidence as follows:

$$720 \quad \begin{aligned} P(g|y) &= \sum_{i=1:4} P(g|y, m_i)P(m_i|y), \\ P(m_i|y) &\propto P(y|m_i)P(m_i), \\ P(y|m_i) &= \int_g P(y|g, m_i)P(g|m_i)dg, \\ P(m_i) &\propto 1, \end{aligned} \quad (3)$$

721 where  $P(g | y)$  is the mixed probability for an individual neuron and  $P(y | m_i)$  is the evidence of

722 the  $i$ th model. Here, all models are treated equally with a non-preference prior  $P(m_i)$ . Finally, the  
723 point estimates of  $g_i$  and  $g_c$  were computed by marginalizing the mixed posterior probabilities.

724

### 725 **Calculation of the synchrony for individual neurons**

726 The spike train of a neuron was binned into  $X(i)$ , where  $i$  represents the time step ( $i = 1, \dots, T$ ),  
727 with  $X(i) = 1$  if the spike occurs in the  $i$ th time bin; otherwise,  $X(i) = 0$ . The synchrony of two  
728 different neurons,  $x$  and  $y$ , was calculated as the cross-correlation coefficient at zero-time lag:

$$729 \quad C_{x,y} = \frac{\sum_{i=1}^T \bar{X}(i)\bar{Y}(i)}{\sqrt{\sum_{i=1}^T \bar{X}(i)^2 \sum_{i=1}^T \bar{Y}(i)^2}} \quad (4)$$

$$730 \quad \bar{X}(i) = X(i) - \frac{1}{T} \sum_{j=1}^T X(j), \quad \bar{Y}(i) = Y(i) - \frac{1}{T} \sum_{j=1}^T Y(j),$$

731 where  $\bar{X}(i)$  and  $\bar{Y}(i)$  are the normalized forms of  $X(i)$  and  $Y(i)$  to account for the firing  
732 frequency. Here, the two spikes were considered synchronous if their onsets occur in the same  
733 1 ms bin. The synchrony level of an individual neuron  $x$  was computed as the mean of  $C_{x,y}$  for  
734 all neurons  $y \neq x$  in the same animal.

735

### 736 **Estimation of the dimensionality of neural firings**

737 The dimensionality can be considered as the minimal dimensions necessary to provide accurate  
738 description of neural dynamics. The principal component analysis (PCA) has become the most  
739 widely used approach because it enables to represent neural dynamics in a lower dimensional  
740 space (Mazzucato et al. 2016). Here, we adopted this approach for estimating the  
741 dimensionality of the IO firing activity in the presence of a small number of recorded neurons.

742

743 We first segmented the IO spike trains into time windows of 50 seconds, from which the firing  
744 rate vectors of all neurons were computed (see Fig S5A). Firing rate vector in each sampled  
745 window corresponds to an observation in the  $N$ -dimensional space, where  $N$  is the number of  
746 ensemble neurons. Then, PCA was applied to estimate the dimensionality as (Abbott et al.  
747 2011):

$$748 \quad d = \frac{1}{\sum_{i=1}^N \tilde{\lambda}_i}, \quad (5)$$

749 where  $\tilde{\lambda}_i = \lambda_i / (\sum_j \lambda_j)$  are the principal eigenvalues expressed as the amount of variance  
750 explained (see Fig S5A), and  $\lambda_i$  is the  $i$ th eigenvalue of the covariance matrix of the firing rate  
751 vectors.

752

753 It is noted that the dimensionality was not sensitive to the length of sampled window (10-50  
754 seconds were analyzed but no significant different values were found, Fig S5B) probably  
755 because the IO firings are rather stable across the time course. However, it has been shown  
756 that dimensionality estimation depends on the number of ensemble neurons  $N$ . Specifically,  $d$  is  
757 underestimated for small  $N$  but becomes independent of  $N$  for sufficiently large  $N$  (Mazzucato et  
758 al., 2016). After data selection (see above), the number of IO neurons in each animal is  $n = 3$ -  
759 22, which is likely to suffer from the under-sampling bias. To overcome such challenge, we  
760 computed the corrected values of dimensionality following the quadratic extrapolation method  
761 (Shew et al., 2011). First, we randomly selected a fraction  $f$  of  $N$  neurons. We then computed  
762 the dimensionality  $d$ , follows the Equation 5, for fractions  $f = 0.2$  to 1 in steps of 0.2, repeated 50  
763 times for each  $f$ . Next, the average  $d$  versus  $f$  data was fitted by the following model (Fig S5C):

$$764 \quad d(f) = d_0 - \frac{a}{fN} - \frac{b}{(fN)^2} \quad (6)$$

765 The fit parameter  $d_0$  is the corrected estimate of the dimensionality and is reported in the text.

766

## 767 **Computation of the complexity entropy**

768 The Lyapunov exponents quantify the sensitivity of a dynamical system to initial conditions, and  
769 thus are often used as indicators of chaos (Farmer and Sidorowich, 1987; Sano and Sawada,  
770 1985). A number of methods have been developed to compute the Lyapunov exponents from  
771 time series with a fixed sampling interval (Kantz, 1994; Rosenstein et al., 1993). Those  
772 methods, however, are not applicable for our IO data because computation of Lyapunov  
773 exponents requires access to continuous variables, which is not the case in our discrete IO  
774 spike sets. We therefore adopted a previously proposed approach (Hirata et al., 2008) that  
775 approximates the Lyapunov exponents via a recurrence plot by using the edit distance of spike  
776 trains (Victor and Purpura, 1997). Our method requires computing the modified edit distance of  
777 the spike trains (Hirata and Aihara, 2009) and its recurrence plot (Eckmann et al., 1987; Marwan  
778 et al., 2007). The complexity entropy (Letellier, 2006) was computed from the distribution of the  
779 length of diagonal lines in the recurrence plot (see Figure S6 for illustration of the complexity  
780 method).

781

782 We first sampled the spikes trains in windows of 20 seconds and computed the edit distance for  
783 all pairs of sampled windows. To resolve the issue of discontinuity induced by the difference in  
784 the number of spikes in two sampled windows, we adopted a modified version of edit distance  
785 computation as in Hirata and Aihara (2009). Briefly, for each sampled window, we took into  
786 account the spikes that occur immediately before and/or after the time window, thus resulting in  
787 four derived windows. We then computed the edit distance for a total of 16 (4x4) derived pairs of  
788 the two sampled windows and temporarily assigned the minimum value as edit distance  
789 between them. The edit distance of two derived windows is defined by a total minimal cost for  
790 converting one window to the other (Victor and Purpura, 1997). Allowed operations include

791 deletion or insertion of events (both cost 1 for each event), and shift of events (cost 20% the  
792 amount of shifting in second for each event). The edit distance for all pairs of sampled windows  
793 of 20 seconds with an interval of 2 seconds constitutes a two-dimensional distance matrix. We  
794 then updated the edit distance matrix by the shortest distance connecting any two sampled  
795 windows – see Figure S6A.

796

797 The recurrence plot is constructed by binarizing the edit distance matrix, with the distance  
798 values smaller than a predefined threshold as 1, and the others else as 0 (Eckmann et al.,  
799 1987). The threshold was determined so that 30% of data points in the distance matrix were 1,  
800 as in (Marwan et al., 2007). Next, we extracted the frequency distribution of the length of the  
801 points 1 that form diagonal lines in the recurrence plot. The Shannon entropy of that distribution  
802 has been shown to be inversely proportional to the largest Lyapunov exponent (Letellier, 2006).  
803 We thus used the inverse of Shannon entropy as a measure of chaos for the spike data.

804

805 To validate that complexity entropy is an indicator of chaos, we generated noise-free simulation  
806 data and computed the correlations between complexity entropy and the Lyapunov indexes (c.f.  
807 Figure S6B-C). Note that this approach is possible for the simulation data because we have  
808 access to the continuous trace of the membrane potential. Specifically, we first removed the  
809 noise in the synaptic inputs, and simulated 500-second spike trains for more than 100  
810 conductance values ( $g_i$  varied in 0–1.0 mS/cm<sup>2</sup> and  $g_c$  in 0–2.0 mS/cm<sup>2</sup>) and estimated the  
811 complexity entropy from the simulated spike trains. Next we computed the Lyapunov exponents  
812 of the IO model by the method of (Wolf et al., 1985), and then extracted the largest component,  
813  $\lambda_1$ , as well as the Lyapunov dimension,  $D_{KY}$ , as these are two direct indicators of chaos (Kaplan  
814 and Yorke, 1970).

815

816 **Statistical Analysis**

817 Unless specifically stated elsewhere, all data is reported as *mean*  $\pm$  *std*. The non-parametric  
818 Kruskal-Wallis one-way analysis of variance was used to test whether data groups of different  
819 sizes originate from the same distribution.

820

821 **FIGURE LEGENDS**

822

823 **Figure 1: Conductance and coupling coefficient estimation in three experimental**  
824 **conditions.** A-B: boxplot with values of  $g_i$  (A) and  $g_c$  (B) estimates for the three experimental  
825 conditions: carbenoxolone (CBX, green box), control (CON, black box) and picrotoxin (PIX, red  
826 box). The color conventions are same for subsequent plots. C: the effective coupling  $g_{eff}$   
827 computed as the equation (1) for the three conditions. D: the coupling coefficient (CC) estimated  
828 for the three conditions via simulations. Each boxplot shows white line as the mean, dark region  
829 as 95% CIs and light region as 1 std. Asterisks represent significance levels: ns  $p > 0.05$ , \* $p <$   
830  $0.05$ , \*\*\*\* $p < 0.0001$ .

831

832 **Figure 2. Examples of inferior olive firing for the model and the data.** Raster plot of ten  
833 representative inferior olive neurons of the model and the experimental complex spike data of  
834 three animals in the three conditions. A. Carbenoxolone (animal #1, repetitive spiking). B.  
835 Control (animal #7, chaotic) C. Picrotoxin (animal #9, highly synchronous spiking).

836

837 **Figure 3: The synchrony and the dimensionality in IO firings moderated by effective**  
838 **coupling.** A-C: The synchrony estimated as coincidence of spikes in 1 ms bins (A) and  
839 estimation of dimensionality of IO firings (C) for 9 animals in the three data conditions. Error  
840 bars are STDs. Asterisks represent significance levels of t-tests: ns  $p > 0.05$ , \* $p < 0.05$ . B-D:  
841 The synchrony (B) and the dimensionality (D) as functions of effective coupling strength  
842 averaged for selected neurons in individual animals confirming that effective coupling is a  
843 control parameter to optimize the synchrony and thus the dimensionality of IO firings. Each type  
844 of symbol represents the data of an individual animal. The cyan solid lines show results of the

845 linear regression models and shaded regions are of 95% CIs.

846

847 **Figure 4: Inverted U-shaped relationship of the complexity versus effective coupling. A-B:**

848 Complexity entropy versus effective coupling. Upper panel: chaotic levels measured by the

849 complexity entropy of the spike data as a function of effective coupling strength for the model

850 (A) and real inferior olive neurons (B) confirming that moderate couplings induce chaos. Each

851 value in the model (open symbols) is of the model neuron that best fits to the actual IO neuron

852 in terms of the PCA error. Each type of symbol represents the data of an individual animal. The

853 cyan solid lines indicate the second-order of linear model (A) and mixed-effects model (B) and

854 shaded regions are of 95% CIs. Lower panel: spike trains of the representative neurons (located

855 at dark arrows in the upper right plot) which show periodic and synchronous firings for either low

856 or high couplings but exhibits chaotic firings for intermediate couplings.

857

858 **SUPPLEMENTAL INFORMATION**

859

860 **Figure S1: Inferior olive firing data set for all animals.** A: Spike data in 50 second of 10  
861 representative neurons in 9 animals with the physiological conditions (CBX and PIX) in the right  
862 and the control condition (CON) in the left columns.

863

864 **Figure S2: Improving the parameter estimates via Bayesian model-averaging.** A: PCA  
865 error rates of the  $g_i$  and  $g_c$  estimates by the segmental Bayesian inference averaged for the  
866 entire IO neurons for CBX, CON, and PIX conditions for four different models (color bars) in  
867 comparison with the previous model (black bar, Onizuka et al., 2013). The error bars are of 95%  
868 CIs. B: Posterior probabilities of a representative IO neuron by individual models and the mixed  
869 posterior probability with the weights determined by the evidence of Bayesian inference.

870

871 **Figure S3: Estimation of the coupling coefficient (CC) by simulation.** We injected a current  
872 pulse of  $-1 \mu\text{A}/\text{cm}^2$  to the cell #5 and recorded the steady-state voltage change of this “master”  
873 cell and its four post-junctional cells (left panel). We computed the CCs for hundreds of  $g_i$  and  
874  $g_c$  values in the range, in which the estimated conductance of the data distributed, and found a  
875 strong positive correlation between the effective coupling and the CC (right panel,  $R^2 = 0.8$ ,  $p <$   
876  $0.0001$ ). It is noted that the non-linear fit represents the nature of deriving  $g_{\text{eff}}$  from  $g_i$  and  $g_c$   
877 follow the Equation (1).

878

879 **Figure S4: Validation of the inverted U-shaped curves.** We investigated whether  
880 intermediate couplings maximize the complexity entropy by applying the two-linear regression  
881 models as follows. We first divided the data into two separate partitions by the intermediate

882 couplings around the mean  $g_{eff} = 0.032$  (black dashed lines). Linear regressions were then  
883 conducted for the two partitions independently for both the model (A) and the data (B). Results  
884 show significant positive coefficients ( $mean \pm sem$ , model:  $2.3 \pm 0.6$ ,  $p = 0.0002$ ; data:  $1.8 \pm 0.8$ ,  
885  $p = 0.03$ ) for the left partitions and significant negative coefficients ( $mean \pm sem$ , model:  $-2 \pm$   
886  $0.4$ ,  $p < 0.0001$ ; data:  $-1.7 \pm 0.6$ ,  $p = 0.009$ ) for the right ones. We further applied a non-  
887 parametric Gaussian Process regression model, which does not assume an explicit relationship  
888 between the coupling and the complexity entropy. Still, we observed inverted U-shaped curves  
889 maximized at around  $g_{eff} = 0.032$  for both the model (C) and the data (D). In sum, these results  
890 support the inverted U-shaped relationship between the effective coupling and complexity  
891 entropy. The right ordinates of C-D represent the first Lyapunov exponents approximated from  
892 the simulation data (c.f Figure S6B), indicating that intermediate couplings induce chaos. The  
893 shaded regions in A-B are of 95% CIs and those in C-D are of  $\pm sem$ .

894

895 **Figure S5. Dimensionality estimation for the spike data of ensemble neurons.** A:  
896 Illustration of the principal component analysis (PCA) for the firing rate vectors extracted from  
897 50-second windows of three neurons of Animal #6 in the CON condition. The estimated  
898 dimensionality  $d = 1.86$  (dashed dark line, c.f Equation 5), indicates that the approximately 2-  
899 dimensional subspace (shaded gray plane) can explain more than 90% of the variance of neural  
900 firing dynamics. B: Estimating dimensionality (Equation 5) with varied window lengths from 10-  
901 50 seconds for 9 animals in the three data conditions showing the robustness of dimensionality  
902 estimation against the window length. C: under sampling correction of dimensionality estimation  
903 for the cases for which the number of recorded neurons is small by quadratic extrapolation  
904 fitting of the dimensionality  $d$  vs. fraction of selected neurons  $f$ .

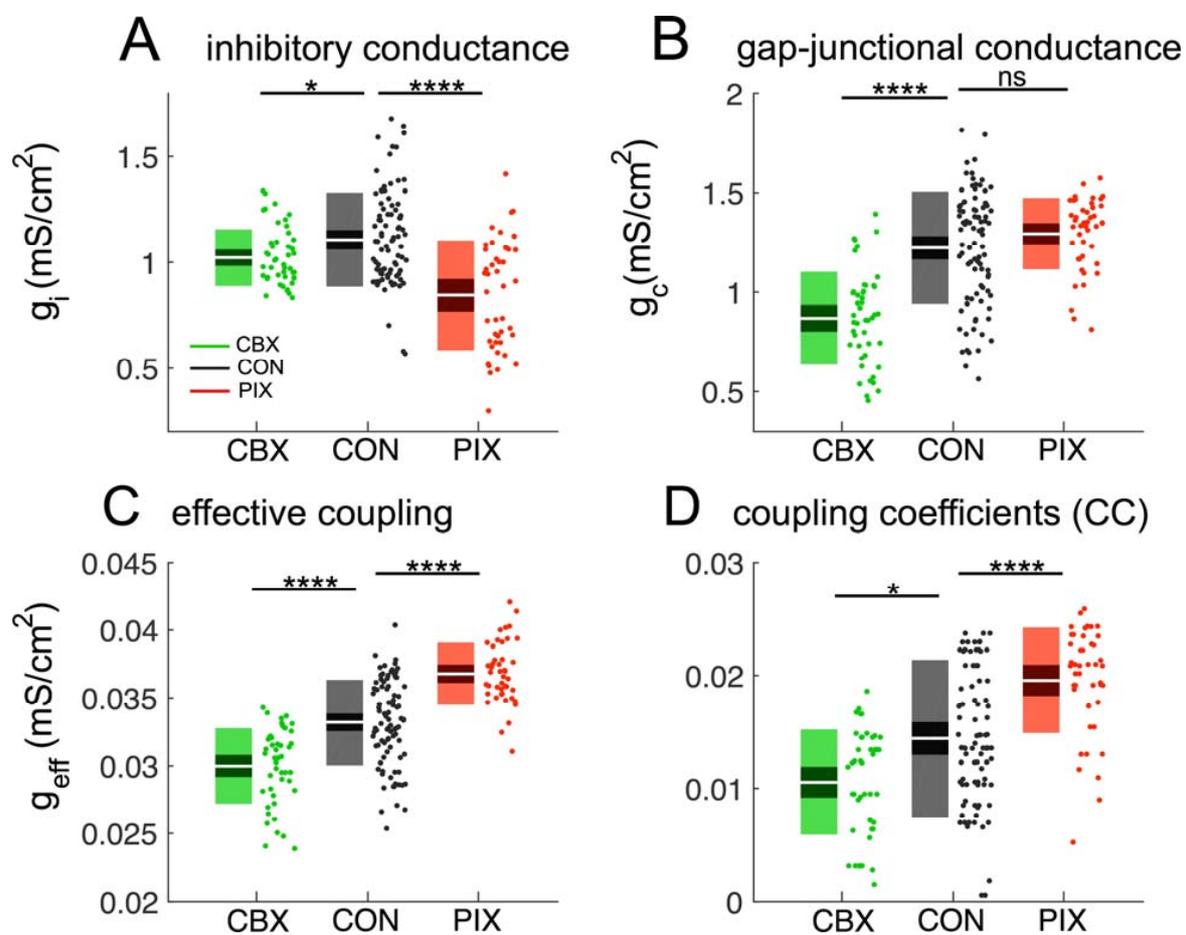
905

906 **Figure S6. Computation and validation of the complexity entropy method.** A: Illustration of  
907 edit distance computation between two sampled spike windows shows a sequence of  
908 elementary steps that convert the spike window (a) into (b). Each bar represents one spike.  
909 Allowed operations include deletion of a spike (shown in red), insertion of a spike (shown in  
910 green), or shifting a spike in time (blue arrows). Computation of edit distance for continuous  
911 sampling windows for the entire spike train constitutes the edit distance matrix. Then, the  
912 recurrent plot is constructed by binarizing the edit distance matrix. The points whose values are  
913 smaller than the threshold were plotted as white dots, otherwise as black dots. Complexity  
914 entropy is computed as the inverse of Shannon entropy, in terms of frequency distribution of the  
915 length of the diagonal lines of white dots (Letellier, 2006). B-C: Complexity entropy measured  
916 for a total of a hundred of parameter values (black crosses) in noise-free simulations showed  
917 strong positive correlations with the largest Lyapunov exponent  $\lambda_1$  (regression model:  $\lambda_1 \sim 1 +$   
918 *entropy*,  $R^2 = 0.4$ , F-test:  $p < 0.0001$ , Figure S6B) and the Lyapunov dimension  $D_{KY}$  ( $D_{KY} \sim 1 +$   
919 *entropy*,  $R^2 = 0.48$ , F-test:  $p < 0.0001$ , Figure S6C). Solid cyan lines represent the fit of linear  
920 models with 95% CIs (dashed cyan lines).

921

922 **Figure 1**

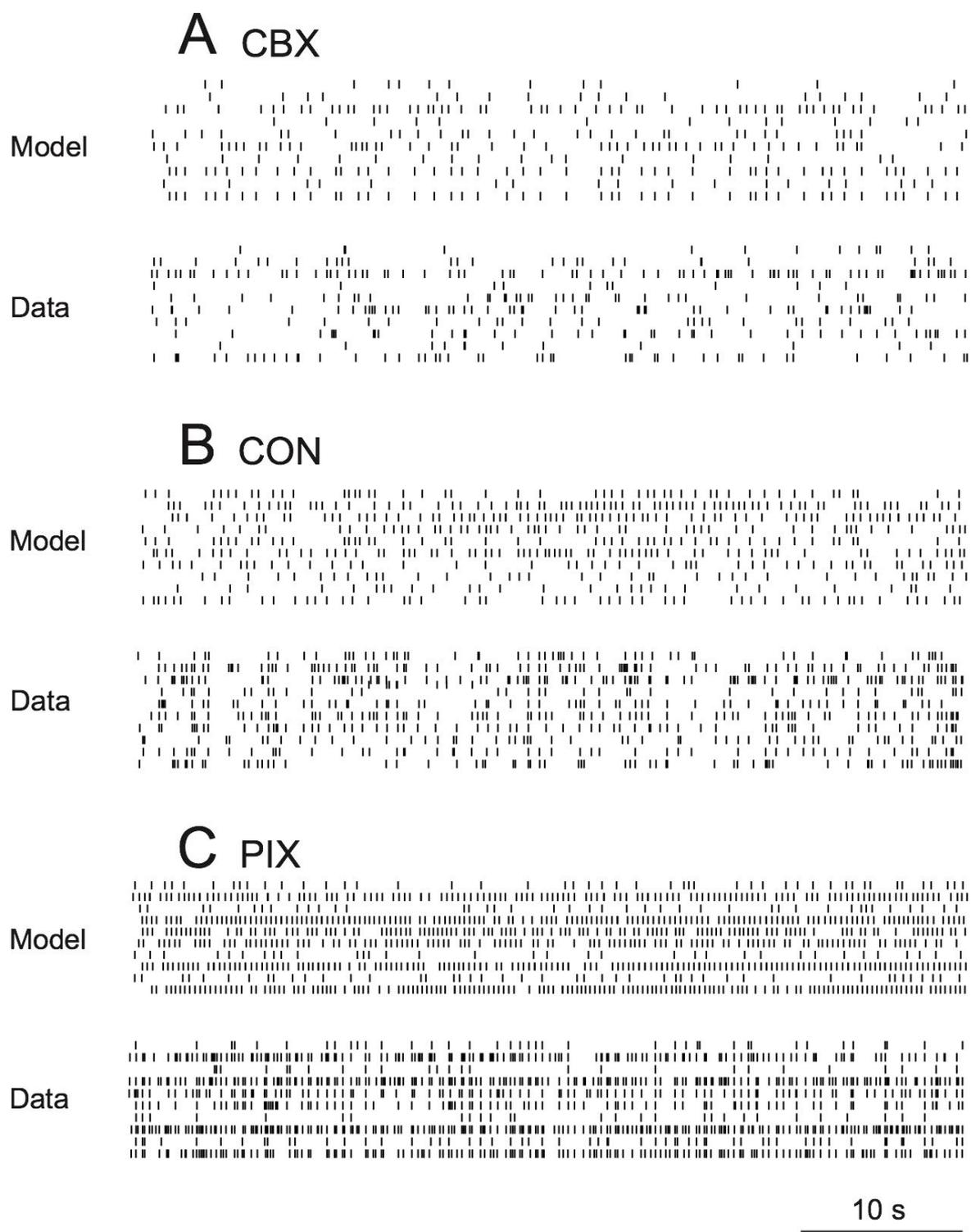
923



924

925

926 **Figure 2**

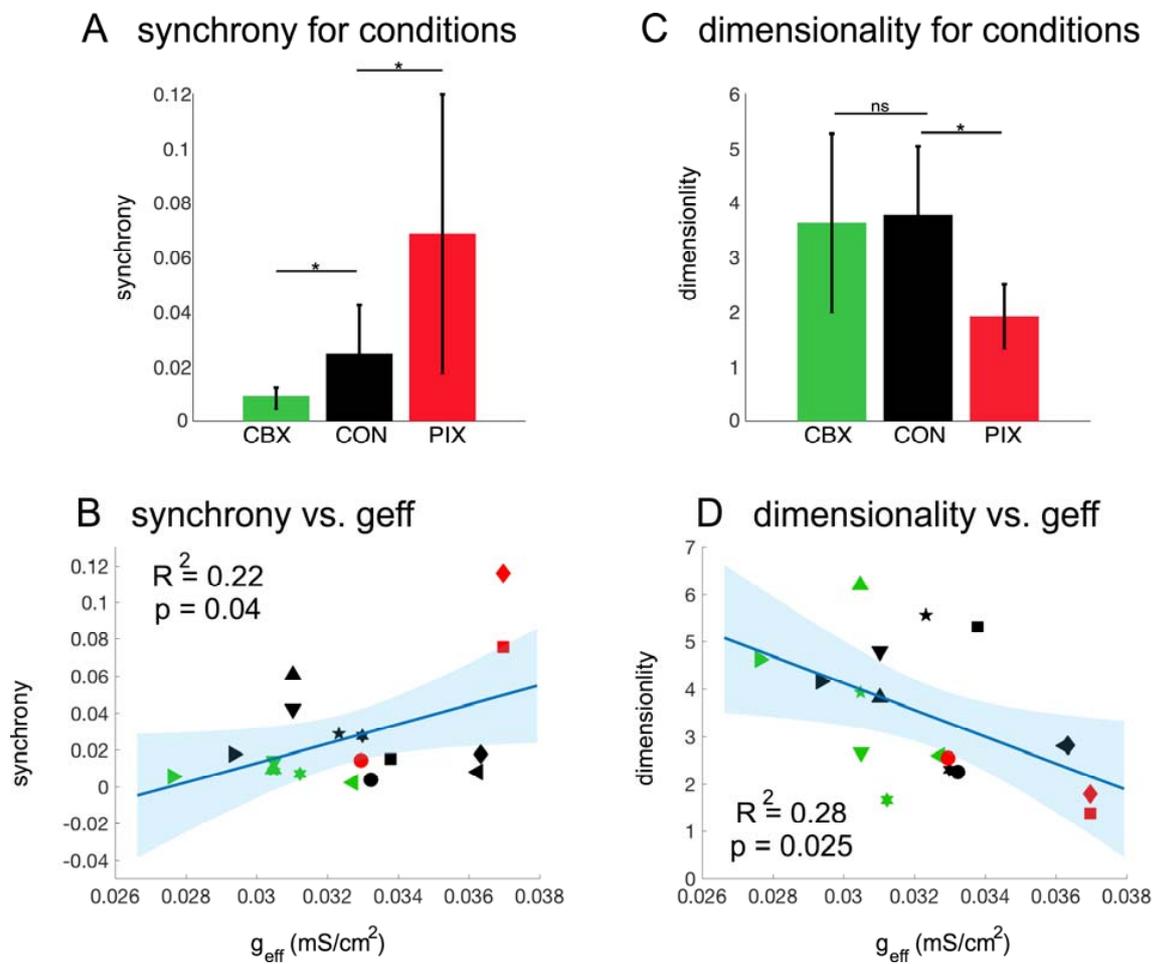


927



929 **Figure 3**

930

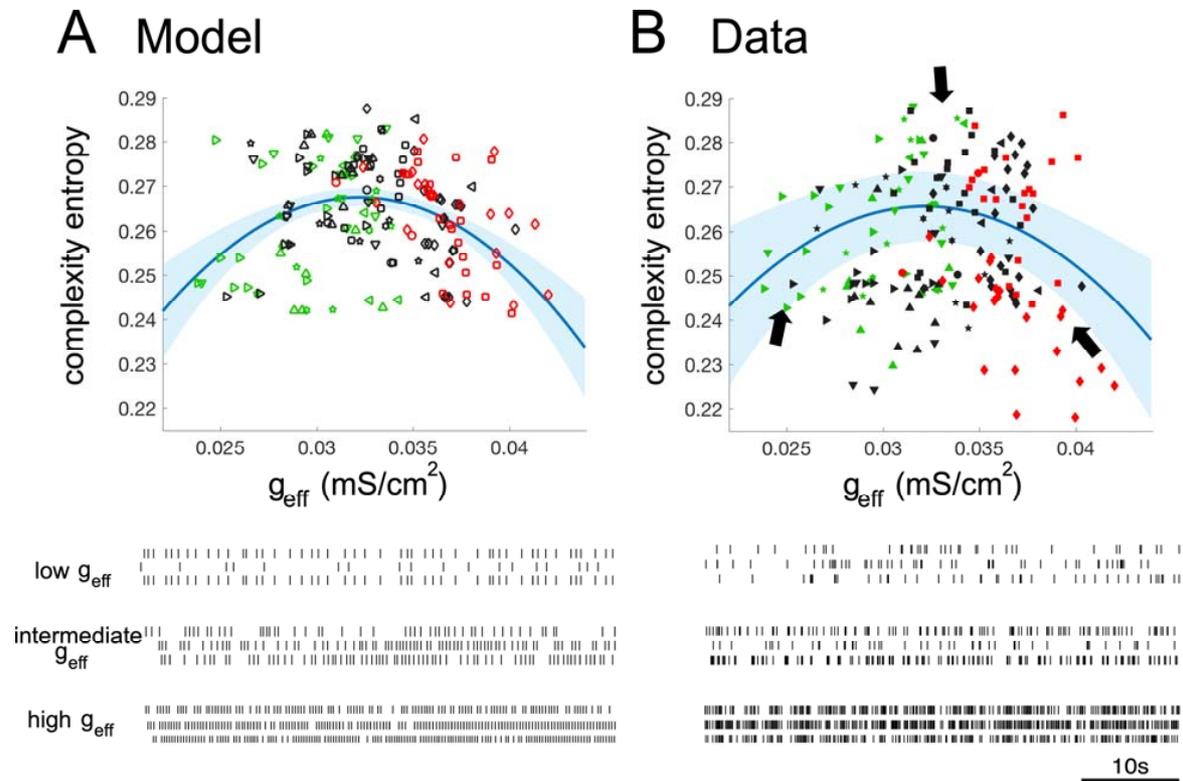


931

932

933 **Figure 4**

934

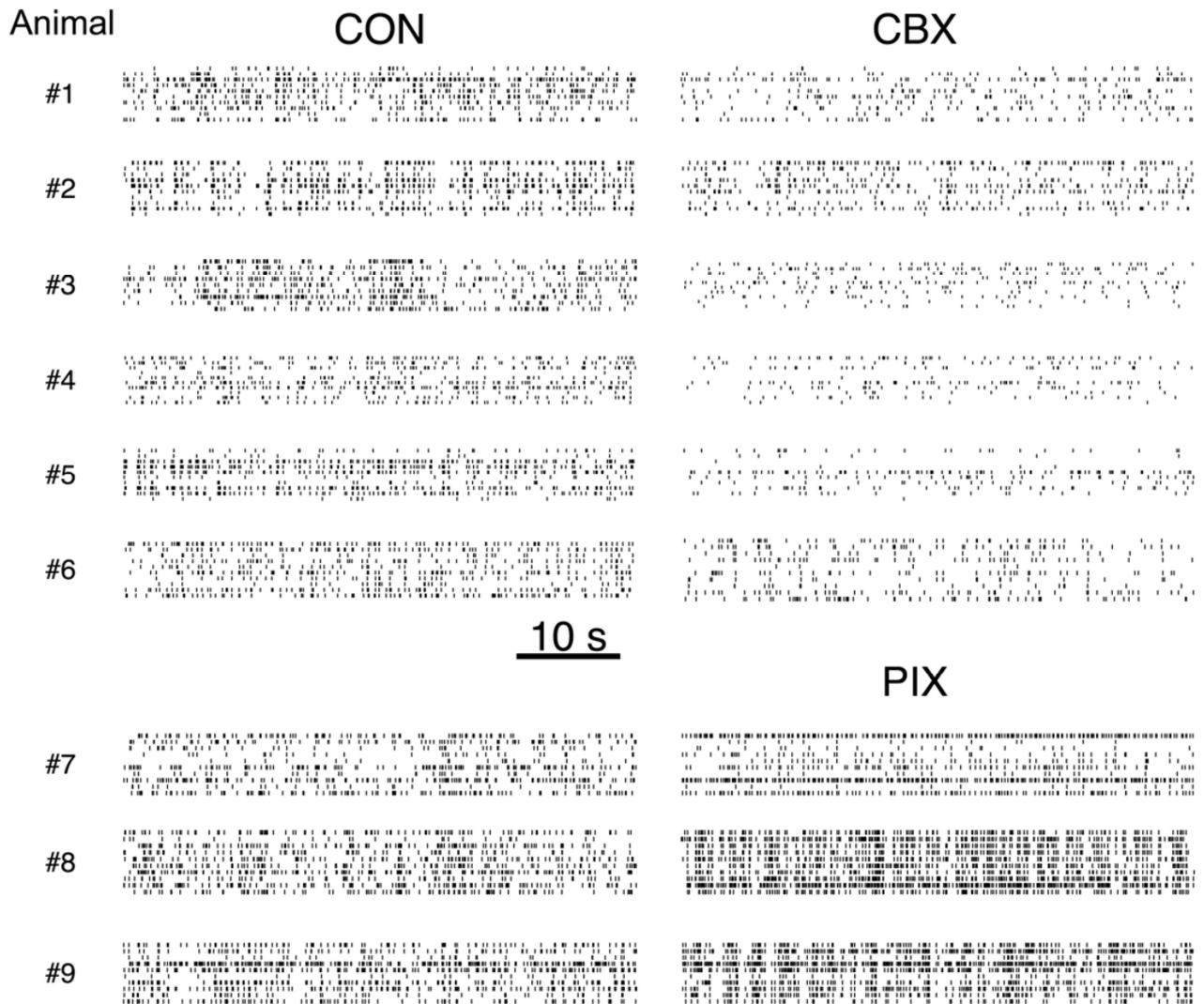


935

936

937 **Figure S1**

938



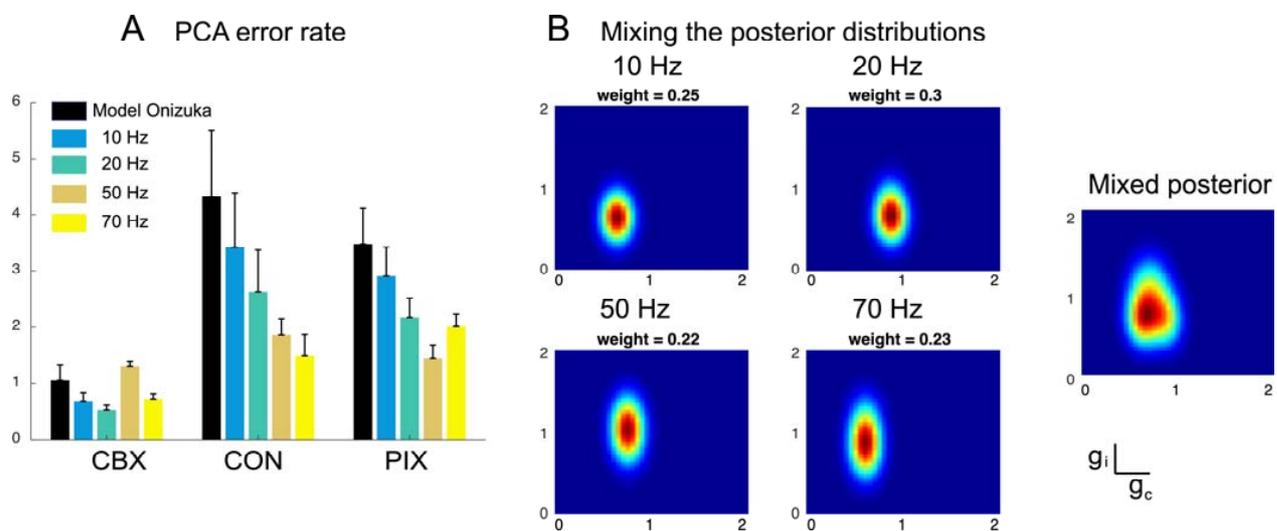
939

940

941

942 **Figure S2**

943

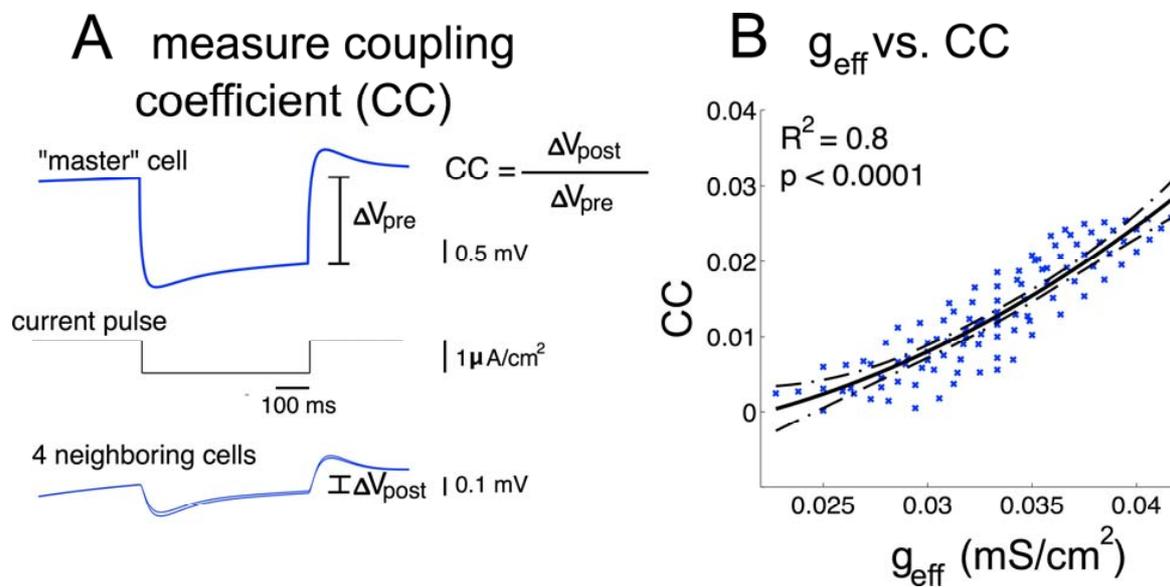


944

945

946 **Figure S3**

947



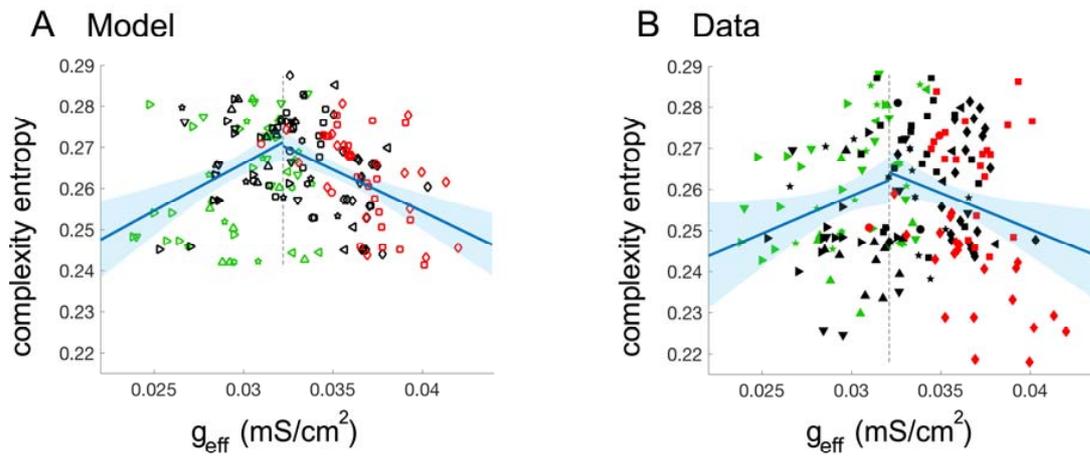
948

949

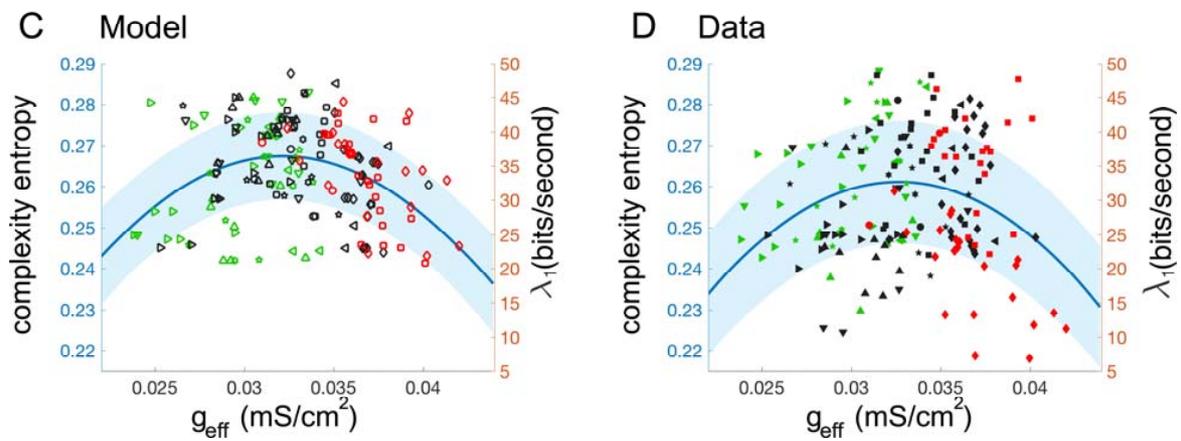
950 **Figure S4**

951

### Two-linear model regression



### Non-parametric Gaussian Process regression



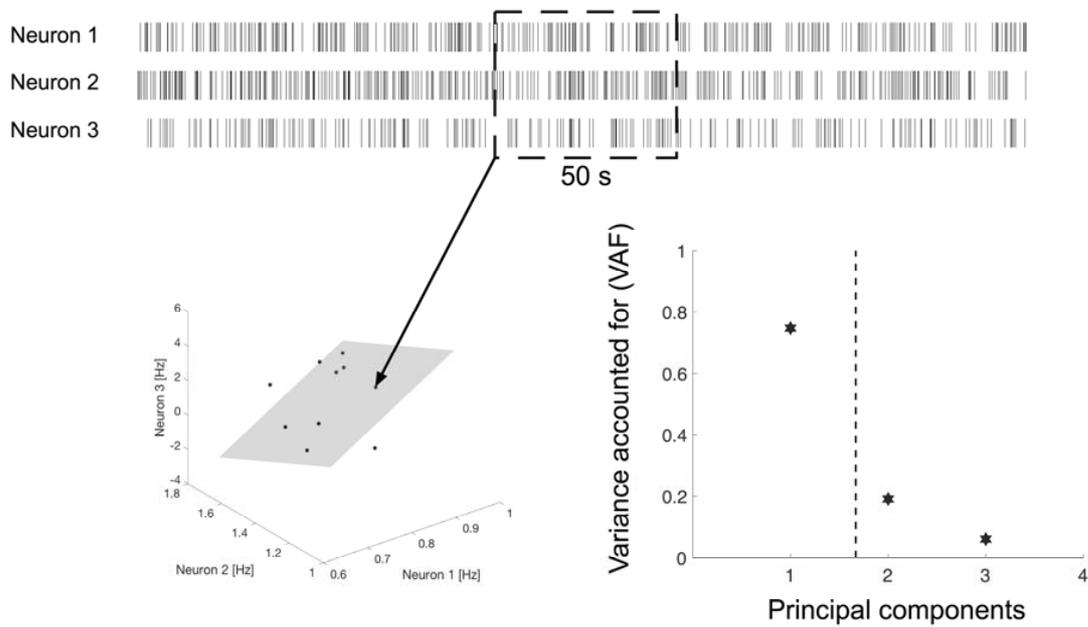
952

953

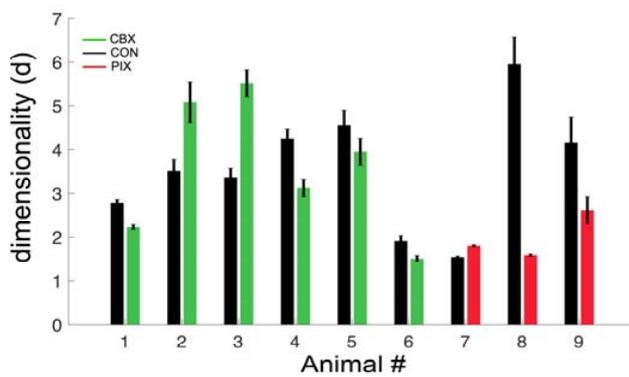
954 **Figure S5**

955

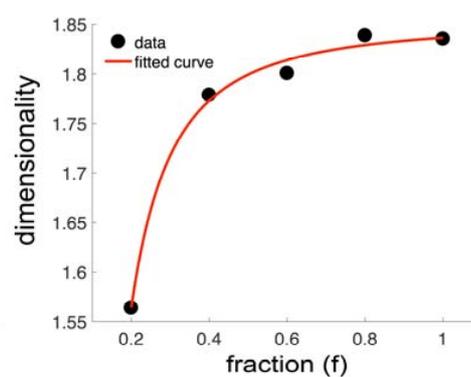
### A dimensionality estimation by PCA



### B dimensionality for varied window sizes



### C correction of estimating d

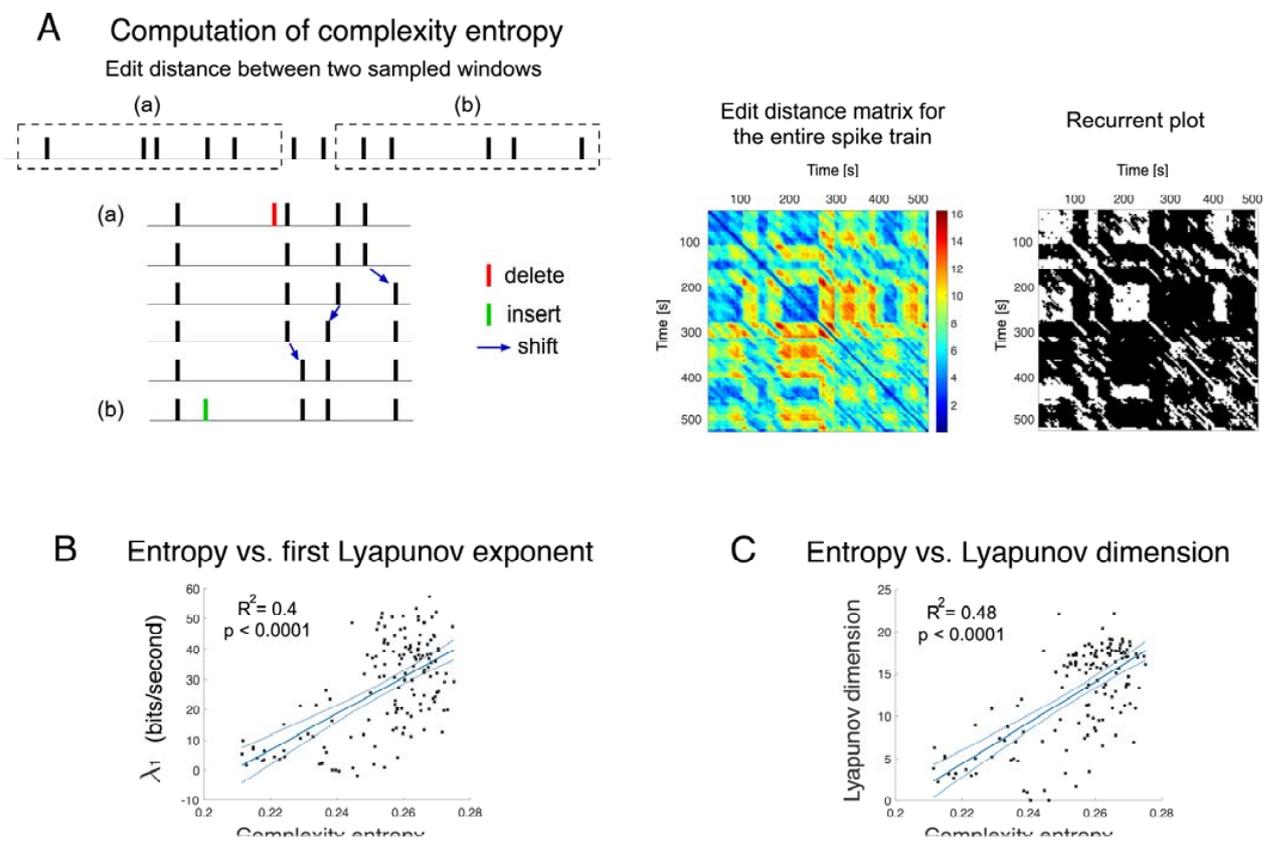


956

957

958 **Figure S6**

959



960