Title: Pattern separation of spiketrains by individual granule cells of the dentate gyrus

Short Title: Temporal pattern separation in the dentate gyrus

Authors: Antoine D. Madar*[1], Laura A. Ewell*[2], Mathew V. Jones[1]

* These authors contributed equally to the manuscript.

Authors affiliations

[1]Department of Neuroscience, University of Wisconsin-Madison, Madison, WI 53705, USA

[2]Laboratory for Experimental Epileptology and Cognition Research, Department of Epileptology, University of Bonn, Bonn, Germany

Corresponding authors

Mathew V. Jones (mathewjones@wisc.edu)

Antoine D. Madar (madar@wisc.edu)

Author contributions

Conceptualization: MVJ, LAE. Data curation: LAE, ADM. Formal analysis: MVJ, ADM. Funding acquisition: MVJ, LAE, ADM. Investigation/data collection: LAE, ADM. Methodology: MVJ, LAE, ADM. Project administration: MVJ, LAE, ADM. Resources: N/A. Software: MVJ, ADM. Supervision: MVJ. Validation: MVJ, LAE, ADM. Visualization: MVJ, LAE, ADM. Writing – original draft: ADM. Writing – review & editing: MVJ, LAE, ADM.

Funding

This work was supported by the University of Wisconsin Institute for Clinical and Translational Research (M.V.J.; NIH/NCATS UL1TR000427) and Lily's Fund for Epilepsy Research (A.D.M.; 2015 fellow).

Abbreviations

GC, granule cell; FS, fast-spiking interneuron; PP, perforant-path; R, Pearson's correlation coefficient; NDP, normalized dot product; SF, scaling factor; SR, spiking reliability; R_w, spiketrain reliability

Abstract

Pattern separation is a process that minimizes overlap between patterns of neuronal activity representing similar experiences. Theoretical work suggests that the dentate gyrus (DG) performs this role for memory processing but a direct demonstration is lacking. One limitation is the difficulty to measure DG inputs and outputs simultaneously. To rigorously assess pattern separation by DG circuitry, we used mouse brain slices to stimulate DG afferents and simultaneously record granule cells (GCs). Output spiketrains of GCs are more dissimilar than their input spiketrains, demonstrating for the first time temporal pattern separation at the level of single neurons in DG. This phenomenon occurs on millisecond to second timescales through different neural codes and is not explained by simple noise. Pattern separation is cell-type specific and larger in GCs than in fast-spiking interneurons. Finally, different GCs process spiketrains differently, a mechanism that likely helps to separate patterns at the population level.

Introduction

How does the brain allow us to discriminate between similar events in our past? This question is a central challenge in the neurobiology of memory and remains elusive. To prevent confusion

between memories that share similar features, the brain needs to store distinct activity patterns to represent distinct memories. In the influential Hebb-Marr framework of episodic memory (1, 2), representations are stored in area CA3 of hippocampus, an auto-associative network where plastic recurrent excitatory connections facilitate recall of stored patterns in response to partial cues (1, 3). However, strong recurrent excitation severely limits the number of patterns that can be stored without overlap (3, 4). Such overlap would lead, when a partial cue common to several patterns is presented, to the reactivation of many patterns and thus to confusion or confabulation. To avoid these interferences, the Hebb-Marr framework proposes that redundancy between input patterns is reduced before they are stored. This process of transforming similar input patterns into less similar output patterns is termed "pattern separation" (4, 5).

Theoretical models suggest that the dentate gyrus (DG) performs pattern separation of cortical inputs before sending its differentiated outputs to CA3 (2, 3). Indeed, DG is ideally located to do this, receiving signals via the major projection from entorhinal cortex (EC), the perforant path (PP), and sending signals to CA3 via granule cells (GCs) axons (6). In addition, behavioral studies have shown that DG lesions impair mnemonic discrimination (7-10). However, although experimental reports have concluded that pattern separation is performed by DG (11-15), they only directly show that similar environments or events are represented differently in the DG. The separation could be done by upstream structures and simply be reported by DG. Hence, it is still unknown whether DG itself performs pattern separation. A rigorous demonstration would require simultaneous knowledge of the inputs arriving at DG and the processed outputs from DG to CA3 (5).

Another difficulty in studying pattern separation is in defining the nature of "activity patterns". Previous studies have focused on spatial patterns of "active neurons", with little

reference to the dynamics of neural activity. For example, computational models predict that DG separates overlapping populations of active EC neurons into less overlapping populations of active GCs (4, 16-19). Immediate-early genes (IEG) expression studies have confirmed that distinct events drive plasticity in different populations of GCs (13, 14, 20) and that overlap in these representations causes mnemonic interference (21). In contrast, *in vivo* single-unit recordings in the DG found that similar contexts are represented by the same population of active neurons, but differences are encoded by different spatially tuned firing patterns (11, 12).

These conflicting results show that pattern separation can correspond to different computations depending on the type of patterns investigated, and that multiple forms of pattern separation could in theory be implemented by DG (5). For example, because *in vivo* recordings suggest that the same neurons are used to code different environments (11, 12), it is possible that pattern separation is performed at the level of single GCs, each disambiguating the activity patterns that it receives. Such disambiguation could be done by changing firing rates, or alternatively, by changing spike timing. Previous experimental investigations of pattern separation in DG examined population vectors of place fields averaged over minutes (11, 12), but place cells also carry information at shorter timescales (22-24). So far, pattern separation has not been well characterized on the scale of milliseconds, and never where patterns are explicitly afferent and efferent trains of action potentials.

Here, we test the hypothesis that DG performs pattern separation of cortical spiketrains, through single GCs, on the millisecond to second timescale. We designed a novel pattern separation assay in acute brain slices to take advantage of the experimental control afforded to slice electrophysiology. Input spiketrains of varying similarities were fed into DG via its afferents, and the output of a GC was simultaneously recorded, allowing the first direct measure

of pattern separation (by comparing input similarity versus output similarity), on timescales relevant to neuronal encoding and synaptic plasticity (23, 25-27).

Results

Temporal pattern separation by individual dentate granule cells

A direct test of pattern separation in single GCs requires knowledge of the similarity between input patterns arriving via the PP, and comparison with the similarity between GC output patterns. Here, we define input and output patterns as rasters of spiketrains. Unless otherwise specified, the similarity between two spiketrains was assessed by computing their pairwise Pearson's correlation coefficient (R) using a binning window τ_w of 10 ms. We generated sets of Poisson input spiketrains (simulating trains of incoming cortical action potentials), with each set having an average correlation R_{input} (**Fig 1A** and **Materials and methods – Pattern separation experiments**). We then recorded the spiking responses of GCs to these sets of input trains delivered to PP fibers (**Fig 1B-C**) (102 recording sets from 28 GCs), allowing us to compute the average output correlation (R_{output}) (**Fig 2A-B**).

For every recording set, R_{output} was lower than the R_{input} of the associated input set, indicating a decorrelation of the output spiketrains compared to their inputs (**Fig 2C**). These results are the first direct experimental evidence that single GCs, the output neurons of DG, exhibit pattern separation. The effective decorrelation, defined as the difference between R_{input} and R_{output}, was statistically significant for every input set, but was larger when input spiketrains were highly correlated (**Fig 2D**). This is consistent with the role of DG in discriminating between similar memories more than already dissimilar ones (8). Note, however, that the decorrelation

normalized to R_{input} is invariant: whatever the input set, the output trains were always decorrelated to about 70% of R_{input} (**Fig 2E**). Such invariance suggests that the same decorrelating mechanism is used on all input sets.

Pearson's correlation coefficient is often used to quantify the similarity between neural activity patterns in computational models (17) and in experimental recordings (11, 12). However, the original Hebb-Marr framework theorized pattern separation as the orthogonalization of the input patterns (1, 5, 28). As a result, the terms "decorrelation" and "orthogonalization" are often conflated in the literature, even though they are not mathematically equivalent and have a non-linear relationship (**S1 Fig** and see **Materials and methods – Similarity metrics**). For instance, pairs of spiketrains can be uncorrelated (R = 0) without being orthogonal, or can be orthogonal without being uncorrelated (**Fig 3A-C** and **S1 Fig.**). To determine whether output spiketrains of GCs are truly orthogonalized, we considered spiketrains as vectors and computed the normalized dot product (NDP) between pairs of spiketrains to assess their similarity (**Fig 3A, C**). For every recording set, NDP_{output} was lower than NDP_{input}, indicating that the angle between output spiketrains was closer to a right angle (i.e., orthogonal) than their inputs (**Fig 3D-E**).

Vectors can differ by their angle, but also by their norm, which in the case of spiketrains is purely dependent on the binwise firing rates. In other words, even if neurons fire in the same time bins (relative to the start of each sweep), the number of spikes per bin can be different, as quantified by the ratio between their norms (scaling factor, SF) (**Fig 3A, C** and **S1 Fig**). Our results show that for very similar inputs, SF_{output} is slightly lower than SF_{input} for most recording sets (**Fig 3F**). This indicates that variations in the binwise firing rate of single GCs in response to similar inputs is a potential, but weak, mechanism of pattern separation at the 10 ms timescale.

As a whole, these results are the first demonstration that input spiketrains are decorrelated in the DG at the level of single GCs, and that this form of pattern separation is mediated by high levels of orthogonalization and weak levels of scaling. As a result, even though R, NDP and SF are not linearly related in theory, R and NDP have a near linear relationship in our dataset, as opposed to R and SF (**S1 Fig E-F**).

Relevant timescales of temporal pattern separation

To measure the similarity of spiketrains we have used metrics that require binning them in time windows of a prespecified size (τ_w). Because the timescales meaningful for the brain remain uncertain, it is important to assess the separation of spiketrains for different τ_w. Our analysis shows that pattern separation, measured through R or NDP, is more pronounced at short timescales (e.g. 5 ms) than at longer ones (≥100 ms) (**Fig 4A-B**). However, although scaling is weak at short timescales it allows DG to perform pattern separation at longer ones (0.5-2 s) through variation of the firing rate (**Fig 4C**).

Because many previous studies suggest that spiketrains can carry information directly through the timing of individual spikes (25-27), we also assessed the similarity between spiketrains using SPIKE, a binless metric purely based on spike times (29). Our results show that input spiketrains with very similar spike times relative to their sweep start (defined here as spiketrains with a high degree of *synchrony*, see **Materials and methods – Similarity metrics**), are transformed into significantly less synchronous outputs, thereby confirming that temporal pattern separation occurs through spike timing modifications in single GCs (**Fig 4D**).

Mechanism of temporal pattern separation

To determine what mechanisms might support temporal pattern separation in GCs, it is necessary to understand its dynamics first. Limiting our analysis to the first presentation of an input set revealed that outputs were already significantly decorrelated (**Figure 5A-B**). This shows that the separation mechanism is fast, consistent with the fact that the brain generally does not have the opportunity to average repeated signals. In addition, analysis of the last presentation revealed only modestly more separation than for the first one, and only for high input correlations (**Figure 5C**), suggesting that learning to recognize the input pattern is not critical.

Because the mechanism for temporal pattern separation is fast and does not require learning, we asked first whether intrinsic properties of GCs could play a role. Linear regression analysis revealed that the membrane capacitance, resistance, time constant as well as the resting membrane potential are not predictors of decorrelation in GCs (see low R^2 in **Table 1**). Another hypothesis is that randomness in neuronal responses drives the decorrelation. Indeed, when the same input spiketrain is repeated (e.g. $R_{input} = 1$) the output spiketrains are not well correlated (as shown by the mean spiketrain reliability R_w) (**Figure 6A**), consistent with well-known trial-to-trial variability in single neuron responses (25, 30, 31). Theoretical investigation of pattern separation often relies on some sort of random process such as probabilistic neuronal activation (4) or stochastic firing (32), which suggests that “neural noise" is a likely contributor to any form of pattern separation. However, because “neural noise” can cover multiple different definitions and phenomena (30), determining its role in a complex computation is not trivial.

Although the noisiness in neural communication is often understood as the unreliability of spiking after a single input spike, and the jitter of the delay between an input spike and an output spike (33), to our knowledge it had never been characterized in GCs before. Hence, we assessed this spike-wise noise in our recordings (**Fig 6B**, **S2 Fig**, and **Materials and methods –**

Noise parameters) and asked whether it can predict the degree of decorrelation by GCs. First, linear regression analysis shows no clear relationship, the spiking reliability (SR) being a mediocre predictor at best (**S3 Fig, Table 2**). Moreover, the average firing rate of a GC output set (a measure directly dependent on SR) is not well correlated with the degree of decorrelation either (**Table 3, S5B Fig**). Thus, even though a relationship might be expected between firing rates and pairwise spiketrains similarity due to higher rates leading to increased probability of spiking close in time (but see **Materials and methods – Similarity metrics**), temporal pattern separation in GCs is not achieved merely because their output spiketrains are sparser than their inputs.

To more carefully test the hypothesis that random spiking failures and delays support fast temporal pattern separation, we produced a shuffled data set and a simulated data set only governed by spike-wise noise statistics comparable to the original data (**S2 Fig** and **Materials and methods – Simulated and shuffled data**). R_{output} was significantly higher in the original data (**Fig 6** and **S4 Fig**), showing that purely random processes yield greater levels of separation than real GCs, especially for highly similar inputs (**Fig 6E** and **S4B Fig)**.

In addition to the spike-wise noise, we considered neural noise at the level of spiketrains using R_w (**Fig 6A**). It characterizes the more complex notion of "spiketrain reliability", that is the ability of a neuron to reproduce the same output spiketrain in response to repetitions of the same input spiketrain. It is not dependent on intrinsic cellular properties (**Table 1**) and only modestly determined by the spike-wise noise (**Table 2**), suggesting that the rather low R_w of GCs is the expression of more complex noisy biophysical processes. Consistently, R_w was significantly lower for shuffled and simulated data than in real GCs (**Fig 6D-E** and **S4C Fig**). This indicates

that the output spiketrains of GCs are more reliable than if their output was entirely determined by simple random processes.

Overall, the lower R_{output} and R_w distributions of random datasets compared to GCs (**Fig 6)** clearly show that simple noise cannot fully underlie the operations performed by GCs on input spiketrains. It also suggests that there might be an unavoidable trade-off between achieving pattern separation and reliable information transmission about input spiketrains. To further test this, we looked at the relationship between R_w and decorrelation levels in individual GC recordings and found a strong anticorrelation (**Fig 7A** and **Table 3**). This is clear evidence that a biological process leading to sweep-to-sweep variability is a powerful mechanism for temporal pattern separation in DG. However, it is not the only source of decorrelation. Indeed, when averaging out the variability between spiketrains associated to the same input, a significant level of decorrelation is still detected (**Fig 7B-E)**. In addition, high levels of pattern separation are achieved in a single sweep (**Fig 5**). This indicates that even if the output spiketrains were perfectly the same from sweep to sweep, they would still be decorrelated compared to their inputs. This makes our discovery that sweep-to-sweep variability is strongly related to temporal pattern separation (**Fig 7A**) even more surprising.

Taken together, these results suggest that complex biophysical mechanisms allow GCs to balance temporal pattern separation and reliable signaling about their inputs.

Fast-spiking interneurons exhibit less temporal pattern separation than GCs

Any brain system might perform either pattern separation or pattern convergence to some degree (5). Thus, GCs are unlikely to be the only neurons to exhibit temporal pattern separation of spiketrains. However, we would expect pattern separation to be at its greatest in GCs, at least

among DG cells, because they are the output neurons of the DG. To test this hypothesis, we performed the same pattern separation assay while recording from fast-spiking interneurons (FS) of the DG instead or in addition to a GC (20 recording sets were collected on 4 FS) (**Fig 8**). We chose FS interneurons because, like GCs, they receive strong input from the PP (34). The distributions of R_{output} were significantly different between the two cell types, with the R_{output} of simultaneously recorded GCs always lower than their corresponding FS (**Fig 8D**). This indicates that FS perform lower levels of decorrelation than GCs.

On the other hand, R_w is significantly higher in FS (**Fig 8E**), illustrating again the trade-off between spiketrain reliability and separation. Surprisingly, in FS, the relationship between R_w and decorrelation follows exactly the same regression line as in GCs (**Fig 7A, Table 3**), suggesting that this trade-off is universal across cell-types.

FS displayed bursting (i.e. more than one output spike between two input spikes) never seen in GCs (34) (**Fig 8C** and **S5A, S5C Fig**) and thus had higher firing rates than GCs (**S5B, D Fig**). Higher firing rates in FS lead to higher correlations between their output spiketrains (**S5B Fig**). We tested whether the bursting was the reason that FS perform less pattern separation than GCs by removing all spikes in a burst except the first one from the FS data set (**S5C Fig**). The resulting dataset ("non-burst" FS: nbFS) had a much lower mean firing rate than FS (**S5D Fig**). However, the degree of correlation of output spiketrains in nbFS was still significantly higher than in GCs (**S5E Fig**). Therefore, bursting and high firing rates are not sufficient to explain the difference between FS and GCs in their ability to separate spiketrains. Interestingly, R_w was also still higher in nbFS than in GCs (**S5F Fig**), showing that the greater ability of FS to reliably transmit information is not a mere consequence of bursting or high firing rates.

Although FS show less pattern separation than GCs, it is interesting that they do exhibit some amount of separation, as opposed to pattern convergence (5) which one could have expected from their reputation of having a much more reliable and precise spiking behavior than principal neurons (33, 35). The high fidelity in relaying input spikes (33) might still explain the difference in pattern separation ability between FS and GCs, although, to our knowledge, they had never been formally compared. We thus first confirmed the idea that FS show much less spike-wise noise than GCs (**S6 Fig**). Then, linear regressions revealed that SR is a good predictor of FS decorrelation performance and R_w (**Table 2**). Surprisingly, the membrane resistance was also a good predictor (**Table 1**). Thus, contrarily to GCs, FS pattern separation behavior is strongly and linearly determined by some intrinsic and spike-wise properties, even though it is in principle hazardous to anticipate complex neuronal operations from such low-level characteristics, as our previous analysis on GCs illustrated.

Overall, these results show that FS perform temporal pattern separation, but to a lower degree than GCs, likely because of their intrinsic properties that allow them to relay information very reliably.

GCs perform temporal pattern separation through distinct operations on input patterns

Pattern separation is typically considered as a population-coding process (5), suggesting that different GCs should take on distinct roles in coding different patterns. To compare GCs to each other, we first assessed the correlation between spiketrains from different cells (not recorded simultaneously) in response to the same input set (**Fig 9A1**). The average correlation $R_{cell\text{-}to\text{-}cell}$ is not dependent on R_{input} and is broadly distributed but skewed towards 0 (**Fig 9A2**). This suggests that different GCs have a general tendency not to fire the same way in response to the same input

spiketrain. On the contrary, FS responses are more consistent between neurons, and FS activity is also poorly correlated with simultaneously recorded GCs (**Fig 9A3**).

Next, we analyzed whether GCs perform pattern separation to the same degree on all pairs of input spiketrains, and compared the amount of decorrelation between different GCs (**Fig 9B**). Individual GCs did not process all input spiketrains in the same way, as demonstrated by the small but significant variability in effective decorrelation for different pairs of input spiketrains (**Fig 9B2**). This variability profile was then used as a fingerprint to be compared across cells that processed the same input set. Our results suggest that different GCs, even from the same animal, can perform pattern separation of the same input spiketrains quite differently. Furthermore, the way pattern separation is performed from GC to GC is more variable and more likely to be different for highly similar input (**Fig 9B3**), which is when pattern separation is theoretically most needed.

These results show that there is variability between different GCs in the way they process and decorrelate input spiketrains, which cannot be attributed to variability in intrinsic cellular properties (**Table 1**).

Discussion

We report that similar input spiketrains are transformed, in GCs, into less similar output spiketrains. Our findings provide the first experimental demonstration that a form of pattern separation is performed within the DG itself and exhibited at the level of single neurons at different timescales through different neural codes. Not all DG neurons perform this computation to the same high degree as GCs, the output neurons of DG to CA3. Finally, temporal pattern

296 separation does not purely result from simple neural noise, but is subject to variability within and
297 between neurons that likely supports pattern separation at the population level.
298
299 **A novel way to test pattern separation**
300 In contrast to *in vivo* experiments that have difficulty identifying recorded units with certainty
301 (36-39) and simultaneously recording the direct inputs of these units (11-15), *in vitro* brain slices
302 that preserve the lamellar connections of the hippocampus offer a more accessible platform. For
303 example, a similar experimental setup to ours was used to show that spatially segregated inputs
304 are represented by distinct spatiotemporal patterns in populations of DG neurons (40, 41).
305 However, our study is the first to perform an experimental analysis of pattern separation within
306 DG by manipulating the similarity of the inputs and comparing it to the similarity of the outputs.
307 Such a systematic approach had so far only been done in computational studies (42). Although a
308 rigorous comparison is impossible because the activity patterns considered were defined
309 differently, the general pattern separation behavior of those models is consistent with our results
310 (**Fig 2C-D**).
311 Studies investigating pattern separation also often differ in the way they measure the
312 similarity between activity patterns. Many methods have been designed to assess similarity
313 between pairs of spiketrains (43-45), each assuming a different definition of similarity. Because
314 we don't know which definition is used by the brain, and given the possibility that multiple ones
315 are relevant, it is important to maintain an agnostic approach. Our study is the first to
316 systematically test pattern separation by considering several similarity measures that span a wide
317 range of potential neural codes (see **Materials and methods – Similarity metrics**). The fact that
318 conceptually different metrics lead to converging results bolsters our conclusion that pattern

319 separation occurs within DG at the level of single GCs. Experiments linking mnemonic
320 discrimination by animals with various potential forms of neural pattern separation will help
321 pinpoint which computations are actually used in episodic memory.
322
323 **Pattern separation through "time" codes.**
324 Until now, most studies of pattern separation in the DG assumed that neural activity patterns
325 were ensembles of ON/OFF neurons (3, 4, 13, 14, 16), sometimes considering a rate code
326 averaged over minutes in addition to this population code (11, 12, 19, 32). Because neurons carry
327 information at timescales shorter than minutes (22-26), and because the sparse firing of active
328 GCs during a brief event (15, 46, 47) precludes an efficient rate code (25), we studied pattern
329 separation at sub-second timescales.
330 Relevant scales are given by the time constant over which neurons can integrate synaptic
331 inputs (23): 10-50 ms for GCs and ~100ms for the "reader" CA3 pyramidal cells. Windows of
332 ~10 ms and ~100 ms, corresponding to gamma and theta rhythms respectively, have been shown
333 to organize CA neuronal assemblies (22, 23, 48, 49). In the DG, spiketrains recorded in similar
334 environments were less synchronous than in CA3 when considering 30-300 ms windows (11). In
335 addition, due to specific network properties allowing persistent activity, the DG might also
336 integrate information over longer time epochs on the order of seconds (40, 41, 50), which have
337 been shown to be relevant in CA fields as well (22). All this suggests that the hippocampus and
338 DG in particular, might convey information through multiple simultaneously relevant timescales.
339 Most of our results are reported at a 10 ms resolution, which corresponds approximately
340 to the spike jitter in GCs (**S2 Fig**) as well as their time constant and the gamma rhythm. This
341 choice of temporal resolution is similar to a recent computational study of pattern separation

342 within a DG model, which used a 20 ms resolution on short spiketrains (30 ms inputs, 200 ms
343 outputs) (17). Furthermore, we found that GCs perform pattern separation both at the millisecond
344 timescale, through orthogonalization or by rearranging spike times, and at the second timescale
345 by varying their firing rate, with a smooth transition around 100 ms (**Fig 4**). Therefore, our work
346 demonstrates for the first time that multiple codes for pattern separation coexist within DG at
347 simultaneously relevant timescales, consistent with a potential multiplexing of signals in the
348 hippocampus.
349
350 **Computational and physiological mechanisms of temporal pattern separation**
351 The mechanisms supporting pattern separation within DG had so far never been experimentally
352 investigated. The orthogonalization of sequentially presented input patterns can in theory be
353 explained by: 1) adaptive mechanisms, involving learning and recognition of input patterns,
354 comparison with previously stored ones and the pruning out of common features, 2) non-
355 adaptive (intrinsic) mechanisms, 3) or both (51). First, concerning adaptive mechanisms, it has
356 been suggested that Hebbian learning could enhance population pattern separation in the DG
357 (52), but computational models testing different forms of synaptic learning found that it would
358 actually impair this type of pattern separation (4, 19). As for temporal pattern separation, our
359 data show that it hardly benefits from the repetition of input patterns (**Fig 5**). Second, we offer
360 indirect evidence that non-adaptive decorrelation processes support temporal pattern separation
361 because output patterns are always decorrelated to the same proportion (**Fig 2E**), a feat that a
362 simple random process can achieve (**Fig S4D**), suggesting that input patterns do not need to be
363 recognized. Third, adaptive and non-adaptive mechanisms are not mutually exclusive: previous
364 learning over days, during the neuron maturation process, could tune single GCs only to specific

365 input patterns, allowing rapid pattern separation (53). Indeed, a computational study suggested
366 that adaptive networks can mature to perform a fast, non-adaptive orthogonalization of the
367 population activity by the decorrelation between individual information channels (54).

368 Adaptive or not, what is the biological source of the temporal decorrelation we observed?
369 Synaptic and intrinsic neural noises are obvious candidates, but simple randomness was not
370 sufficient to reproduce our results (**Fig 6** and **S4 Fig**). More complex and realistic noisy
371 processes including synaptic short-term plasticity as well as inhibition might have a role.
372 However we showed that FS, which provide both feedforward and feedback inhibition to the
373 soma of GCs (34, 35), exhibit poor ability to separate spiketrains. On the other hand, their ability
374 to relay information reliably (35) (**Fig 8** and **S6 Fig**) and to precisely control spike timing in
375 target neurons (35) might actually provide a mechanism that counteracts noisiness in GCs,
376 increasing the fidelity of information transmission to CA3 (while still allowing effective
377 spiketrain separation in GCs).

378

379 **The role of sweep-to-sweep variability**

380 Because the brain needs to be able to recognize when situations are exactly the same, our finding
381 that pattern separation occurs even when the same input pattern is repeated (**Fig 6A**) might seem
382 counter-intuitive at first. However, in theory, the separation and the recognition functions do not
383 have to be supported by the same network. The Hebb-Marr framework actually hypothesizes that
384 CA3 is able to recall the original pattern from a noisy input from DG. Even though most
385 computational models that tested the effect of repetition were consistent with the intuitive view
386 (4, 17), this was likely because they used deterministic neurons. A model considering variability
387 across GCs and a probabilistic spiking behavior had results similar to ours (32).

In the cortex, the well-known variability of single neuron activity between trials is often supposed to be "averaged out" at the population level so that the output of the population is reliable (30). It is thus conceivable that considering an ensemble of GCs would increase the signal-to-noise ratio. In fact, when we average out the sweep-to-sweep variability, GCs exhibit pattern separation for highly similar patterns but almost no separation for identical ones (**Fig 7D-E**).

However, this variability is not necessarily meaningless (30). Our results suggest it might be a mechanism amplifying pattern separation (**Fig 7**). The variability might even be just apparent, if we consider that when the same input is repeated it is at different points in time: each repetition could be considered as a different event that need to be encoded slightly differently. The role of single GCs could thus be to meaningfully add some noise to transform input spiketrains so that cortical information about an event is stored in the hippocampus with a unique random time-stamp, consistent with the index theory of episodic memory (55).

The computational importance of temporal pattern separation in single cells to the population level

Although more work is needed to test whether the DG is a pattern separator at the population level, the discovery of temporal pattern separation in single GCs has strong implications for population dynamics. The fact that, in response to the same patterns, single GCs rearrange their spikes differently from sweep to sweep (**Fig 6A**) and from cell to cell (**Fig 9A**) may enforce very small neuronal assemblies in the DG (23). In other words, these processes may insure that a minimal number of output neurons are active at the same time: such sparsity in active neuronal

410 population is known, from computational studies, to be critical for efficient population pattern
411 separation (4, 18, 56).
412
413 # Materials and Methods
414
415 ## Animals and dissection
416 Horizontal slices (57) of the ventral and intermediate hippocampus (400 μm) were prepared from
417 the brains of C57BL/6 male mice 15 – 25 days old (Harlan). All procedures were approved by
418 the University of Wisconsin Institutional Animal Care and Use Committee. Mice were
419 anesthetized with isoflurane, decapitated, and the brain was removed quickly and placed in ice-
420 cold cutting solution containing (in mM) 83 NaCl, 26 $NaHCO_3$, 2.5 KCl, 1 NaH2PO_4, 0.5 $CaCl_2$,
421 3.3 MgCl2, 22 D-Glucose and 72 Sucrose, bubbled with 95% O_2 and 5% CO_2. Horizontal slices
422 were cut using a vibratome (Leica VT1000S) and placed in an incubation chamber in standard
423 artificial cerebrospinal fluid (aCSF) containing (in mM) 125 NaCl, 25 NaHCO3, 2.5 KCl, 1.25
424 NaH2PO4, 2 CaCl2, 1 MgCl2, and 25 D-Glucose (or in a 50/50 mix of cutting solution and
425 standard aCSF) at 35° C, for 15-30 minutes after dissection. Slices were stored in the incubation
426 chamber at room temperature for at least 30 minutes before being used for recordings.
427
428 ## Electrophysiology
429 All recordings were done in aCSF. Whole cell patch-clamp recordings were made using an
430 upright microscope (Axioskop FS2, Zeiss, Oberkochen, Germany) with infra-red differential
431 interference contrast optics. Patch pipettes pulled from thin-walled borosilicate glass (World
432 Precision Instruments, Sarasota, FL) had a resistance of 3-5 MΩ when filled with intracellular

433 solution containing (in mM) 140 K-gluconate, 10 EGTA, 10 HEPES, 20 phosphocreatine, 2
434 Mg_2ATP, 0.3 NaGTP (pH 7.3, 310 mOsm). Recordings were done at physiological temperature
435 (33-35 °C) using one or two Axopatch 200B amplifiers (Axon Instruments, Foster City, CA),
436 filtered at 5 kHz using a 4-pole Bessel filter and digitized at 10 kHz using a Digidata 1320A
437 analog-digital interface (Axon Instruments). Data were acquired to a Macintosh G4 (Apple
438 Computer, Cupertino, CA) using Axograph X v1.0.7 (AxographX.com). Stimulation pipettes
439 were pulled from double barrel borosillicate theta-glass (~10 μm tip diameter, Harvard
440 Apparatus, Edenbridge, U.K.) and filled with ACSF or a 1M NaCl solution and connected to a
441 constant current stimulus isolator used to generate 0.1-10 mA pulses, 100 microseconds in
442 duration. GCs used for analysis (n = 28) were stable across a whole recording session as judged
443 by monitoring of series resistance and resting potential, with the following characteristics: series
444 resistance (R_s): 6.65 ± 0.68 MΩ; resting potential (V_{rest}): -69.3 ± 1.3 mV (min = -80 mV, max = -
445 51 mV); input resistance (R_i): 171 ± 16MΩ (min = 81 MΩ, max = 325 MΩ) and capacitance
446 (C_m): 23 ± 2 pF (min = 12 pF, max = 65 pF). Fast-spiking (FS) interneurons (n = 4) were
447 identified as neurons with large somata at the hilus-granule cell layer border and a high firing
448 rate response during large depolarizing current steps (34, 58) (Fig. 7). They had the following
449 characteristics: R_s: 7.2 ± 1.2 MΩ; V_{rest}: -66.7 ± 3.5 mV (min = -72 mV, max = -55 mV); R_i: 59 ±
450 10 MΩ (min = 41 MΩ, max = 92 MΩ) and C_m: 19 ± 3 pF (min = 13 pF, max = 30 pF).
451
452 **Pattern separation experiments**
453 Input patterns were 2 second long traces of impulses simulating cortical spiketrains, with
454 interspike intervals following a Poisson distribution of mean frequency ~10 Hz (11.9 ± 0.7 Hz,
455 min = 9.6 Hz, max = 14.5 Hz). Firing rates were chosen to be consistent with the frequency of

EPSCs recorded in GCs of behaving mice (46), and are known to promote a high probability of spiking in GCs in slices (34, 59). Eleven sets of five input trains were designed so that the 5 trains of each set would have a prespecified average correlation coefficient R_{input} when using a binning window τ_w of 10 ms. The relative standard error of the input set similarity was on average 4% of the mean for R_{input} at τ_w = 10 ms, and it was similarly constrained for other time resolutions and similarity metrics. Five sets were designed with an algorithm developed in-house (at τ_w = 10 ms, R_{input} = 0.88, 0.84, 0.73, 0.65, 0.56) and six other sets were designed using the algorithm of Macke and colleagues (60) (at τ_w = 10 ms, R_{input} = 1.00, 0.95, 0.76, 0.48, 0.25, 0.11). Because results did not qualitatively differ when considering data obtained from the two groups of input sets, we pooled them together for our analysis.

The spiking response of a DG neuron was recorded in whole-cell mode while stimuli were delivered to the outer molecular layer (OML). Stimulus current intensity and location were set so that the recorded neuron spiked occasionally in response to electrical impulses and the stimulation electrode was at least 100 µm away from the expected location of the dendrites of the recorded neuron. Once stimulation parameters were set, a pattern separation protocol was run. It consisted of a sequence of the five different input spiketrains, delivered one after the other separated by 4 s of relaxation, repeated ten times. The ten repetitions of the sequence of five patterns were implemented to take into account any potential variability in the output, and the non-random sequential scheme was used to avoid repeating the same input spiketrain close in time. Each protocol yielded a recording set consisting of fifty output spiketrains, each associated with one of the five different input spiketrains (**Fig 1C**). V_{rest} was maintained around -70mV during recordings, consistent with the V_{rest} of mature GCs recorded in behaving mice (46). The output spiking frequency was variable (6.3 ± 0.3Hz, see **S5B Fig**) but consistent with sparse

activity generally observed in GCs *in vivo* (11, 12, 46, 47, 61) and in slices under conditions of drive comparable to what was used here (62, 63).

Similarity metrics

Similarity between spiketrains was assessed in four ways: 1) with the Pearson's correlation coefficient (R), 2) with the normalized dot product (NDP), 3) with the scaling factor (SF) and 4) with a distance metric called SPIKE specifically designed to assess the dissimilarity between two spiketrains (29). The SPIKE-metric is a binless metric based on spike times, whereas R, NDP and SF are based on the number of spikes occurring in time bins of prespecified durations (e.g. τ_w = 10 ms).

For R, two spiketrains X and Y of the same duration, divided into N time bins of size τ_w are seen as variables, with X_i and Y_i the observations, i.e. the respective numbers of spikes in bin i for each spiketrain. R assesses the similarity between X and Y by measuring the goodness of fit of a linear regression to the distribution of points (X_i, Y_i): when R is close to 1 spiketrains are similar, close to 0 they are dissimilar and close to -1 they are anticorrelated (**Fig 3A-B)**. R was computed with the following equation, where *cov* is the covariance, σ is the standard deviation and $\bar{X}$ and $\bar{Y}$ are the means of X and Y):

$$R = \frac{cov(X,Y)}{\sigma_X . \sigma_Y} = \frac{\sum_{i=1}^{N}(X_i - \bar{X}).(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{N}(X_i - \bar{X})^2} . \sqrt{\sum_{i=1}^{N}(Y_i - \bar{Y})^2}}$$

NDP and SF are similarity metrics explicitly considering spiketrains as vectors. They, like R, require arbitrarily dividing spiketrains into time bins, which are considered dimensions of

an N-dimensional space where N is the number of bins. For two binned spiketrains X and Y, NDP is the cosine of the angle θ between the two vectors: 0 when they are perfectly orthogonal, 1 when they are collinear (**Fig 3A, C**). The NDP is defined as the scalar product of X and Y divided by their norms, and was computed with the following equation (where X_i and Y_i are the coordinates of X and Y, measuring the number of spikes in bin *i*):

$$\mathrm{NDP} = \cos(\theta) = \frac{\sum_{i=1}^{N} X_i . Y_i}{\sqrt{\sum_{i=1}^{N} X_i^2} . \sqrt{\sum_{i=1}^{N} Y_i^2}}$$

SF, on the other hand, quantifies the difference of length between the two vectors X and Y, or, in other words, the variation in the binwise firing rate between two spiketrains. We have defined it as the ratio between the norms of each vector, the smaller norm always divided by the bigger one to have SF values ranging from 0 to 1. When norms ||X||, ||Y|| or both were 0 (i.e. spiketrains without spikes), SF was excluded from further analysis. SF = 1 means X and Y are identical in terms of binwise spike number. The closer to 0 SF is, the more dissimilar are the binwise firing rates (**Fig 3A, C**). SF was computed with the following equation (where $0 < ||X|| \leq ||Y||$):

$$SF = \frac{\sqrt{\sum_{i=1}^{N} X_i^2}}{\sqrt{\sum_{i=1}^{N} Y_i^2}}$$

Comparing the three equations above, the relationships between R, NDP and SF are not trivial and not linear. Note that R is actually an NDP, but of vectors centered to their respective

519 mean, which does not conserve the angular relationship between X and Y. This centering also
520 makes R a similarity metric intrinsically independent of differences in mean firing rates between
521 X and Y (43, 64), as opposed to NDP and SF. (NB: this does not prevent a physiological
522 dependency, as shown by de la Rocha and colleagues (2007) and our results in **S5B Fig**). To
523 further evaluate the relationship between R, NDP and SF, we generated a set of 1012 spike
524 rasters (1000 were random, 12 were specific cases) between which we computed the similarity
525 with the three metrics above. Each raster was made of six bins, each bin containing 0, 1 or 2
526 spikes (drawn from a uniform distribution for the randomly generated group). This analysis
527 confirmed that the three metrics are not equivalent and provides an intuition on what each metric
528 represents (**S1D Fig**).

529 As explained above, R, NDP and SF are binned measures, with τ_w the specified temporal
530 resolution. In other words, in all the equations above, X_i and Y_i are functions of τ_w and the values
531 of the respective similarity metrics are dependent on τ_w as well. In **Fig 4**, we evaluated the
532 influence of τ_w on the results by varying it between 5 ms and 2000 ms. Note that because our
533 spiketrains are 2 seconds long, using $\tau_w = 2000$ ms means the spiketrains can be seen as variables
534 with only one observation, or as unidimensional vectors that can only vary by their norm. In this
535 case, R is meaningless, NDP is necessarily 1, which indicates collinearity, and SF correctly
536 assesses the variation in the overall firing rate between the two spiketrains. Our analysis
537 therefore explores a wide range of coding strategies between a temporal code with 5 ms
538 resolution and a pure rate code.

539 Being binned metrics, R, NDP and SF also have the drawback of considering all X_i (i.e.
540 bins) as independent observations, which may not be a realistic assumption. The binless SPIKE
541 similarity metric avoids this limitation. SPIKE also differs from the other measures by not

assuming that spiketrains are related linearly, or that they belong to a Euclidean space (65). To compute SPIKE, we used the Matlab toolbox provided at **www.fi.isc.cnr.it/users/thomas.kreuz/sourcecode.html** which computes the SPIKE-distance (called D(t) in our study and S(t) in the original paper: see equation 19) (29). The SPIKE similarity was computed as: SPIKE = $1 - \frac{1}{T}\int_0^T D(t)dt$, where T is the duration of the spiketrain. Because D(t) ranges from 0 to 1, SPIKE is thus also between 0 and 1, like NDP and SF. When SPIKE is equal to 1, spiketrains have exactly the same spike times, i.e. they are synchronous (n.b. in our experiments, spiketrains were not simultaneously recorded, but we use "synchronous" in the sense of spiketrains aligned to the start of each 2 s sweep). Note that SPIKE has a large dynamic range (i.e. sensitivity over large differences of spiketimes), and, as a result, realistic spiketrains like in our input sets rarely have a SPIKE similarity lower than 0.5 (29) (**Fig 4C**).

$-1 \leq R \leq 1$
$0 \leq NDP \leq 1$
$0 < SF \leq 1$
$0 \leq$ SPIKE-similarity ≤ 1; but ≥ 0.5 in most cases

Noise parameters. We define the spike-wise neural noise as the delay of an output spike after an input spike, its average jitter and its spiking reliability (SR, below) which is linked to the rate of failure to spike after an input spike. To assess these parameters, we computed the cross-occurrence between input spikes and output spikes in a [-15 ms, 50ms] interval with 1 ms bins, for each recording set. The resulting histogram of counts of output spikes occurring in the vicinity of an input spike was fitted with a Gaussian distribution $N(\mu,\sigma$, baseline), where μ is the

mean delay of an output spike and σ is the jitter of this delay. The baseline corresponds to the background firing, occurring by chance or caused by neighboring inputs. After subtracting the baseline and extracting the probability of spiking by dividing the counts of output spikes by the total number of input spikes, we defined the *spiking reliability* (SR) as the sum of probabilities of an output spike in the predefined time interval around an input spike (**Fig 6B** and **S2A Fig**).

Simulated and shuffled data. To assess the role of spike-wise neural noise in pattern separation, we generated two data sets. First, we simulated output spiketrains in response to our 11 input sets (10 simulated output sets of 50 synthetic spiketrains per input set). This simulation was entirely based on the average spike-wise noise parameters computed from the real GC recordings (see above): the matrix of input spike times was replicated ten times, and for each of the 50 resulting sweeps, spikes were deleted randomly following a binomial distribution B(Number of spikes, 1-mean SR = 1-0.42). A random delay, sampled from a Gaussian distribution $N(\mu,\sigma)$, was added to each resulting spike times, with μ and σ being respectively the mean delay and mean jitter in the original recordings. The noise statistics of the resulting simulated data set is shown in **S2C Fig.**

Second, we created a surrogate data set by randomly shuffling the output spikes of the original GC recordings: the delay of each spike was conserved but it was relocated to follow a randomly selected input spike in the same input train (from a uniform distribution). This strategy yielded a data set with noise statistics closer to the original data (**S2D Fig)**.

Software and statistics

Data analysis was performed using custom-written routines in Matlab (Mathworks, Natick, MA, USA), including functions from toolboxes cited above. Sample sizes were chosen based on the

584 literature and estimations of the variance and effect size from preliminary data. All values are
585 reported as mean ± S.E.M. unless otherwise noted. The one-sample Kolmogorov-Smirnov test
586 was used to verify the normality of data distributions. Parametric or non-parametric statistical
587 tests (see figure legends) were appropriately used to assess significance (p-value < 0.05).
588 Assumptions on equal variances between groups were avoided when necessary. All T and U tests
589 were two-tailed. To determine whether two distributions of data points are significantly different
590 (e.g. R_{output} as a function of R_{input}, for GC compared to FS, see **Fig. 5, 6, S4, 7, S5**), we performed
591 a regression (linear or parabolic) on the two data sets as well as on the combined data set, and
592 assessed significance via an F-test comparing the goodness of fits (66). Because R_{input} can also be
593 considered as a categorical variable, we performed a two-way ANOVA before using post-hoc
594 tests correcting for multiple comparisons in order to determine at which R_{input} groups two
595 conditions were significantly different.
596

597 ## References

598 1. McNaughton BL, Morris RGM. Hippocampal synaptic enhancement and information
599 storage within a distributed memory system. Trends in Neuroscience. 1987;10(10):408-15.
600 2. Treves A, Tashiro A, Witter MP, Moser EI. What is the mammalian dentate gyrus good
601 for? Neuroscience. 2008 Jul 17;154(4):1155-72. PubMed PMID: 18554812. Epub 2008/06/17.
602 eng.
603 3. Rolls ET. A computational theory of episodic memory formation in the hippocampus.
604 Behavioural brain research. 2010 Dec 31;215(2):180-96. PubMed PMID: 20307583.

4. O'Reilly RC, McClelland JL. Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. Hippocampus. 1994 Dec;4(6):661-82. PubMed PMID: 7704110. Epub 1994/12/01. eng.

5. Santoro A. Reassessing pattern separation in the dentate gyrus. Frontiers in behavioral neuroscience. 2013;7:96. PubMed PMID: 23908611. Pubmed Central PMCID: 3726960.

6. Amaral DG, Scharfman HE, Lavenex P. The dentate gyrus: fundamental neuroanatomical organization (dentate gyrus for dummies). 2007;163:3-790.

7. McHugh TJ, Jones MW, Quinn JJ, Balthasar N, Coppari R, Elmquist JK, et al. Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. Science. 2007 Jul 6;317(5834):94-9. PubMed PMID: 17556551.

8. Kesner RP, Rolls ET. A computational theory of hippocampal function, and tests of the theory: new developments. Neuroscience and biobehavioral reviews. 2015 Jan;48:92-147. PubMed PMID: 25446947. Epub 2014/12/03. Eng.

9. Kesner RP, Kirk RA, Yu Z, Polansky C, Musso ND. Dentate gyrus supports slope recognition memory, shades of grey-context pattern separation and recognition memory, and CA3 supports pattern completion for object memory. Neurobiology of learning and memory. 2016 Mar;129:29-37. PubMed PMID: 26318932. Epub 2015/09/01. eng.

10. Baker S, Vieweg P, Gao F, Gilboa A, Wolbers T, Black SE, et al. The Human Dentate Gyrus Plays a Necessary Role in Discriminating New Memories. Current biology : CB. 2016 Sep 21. PubMed PMID: 27666968. Epub 2016/09/27. Eng.

11. Leutgeb JK, Leutgeb S, Moser MB, Moser EI. Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science. 2007 Feb 16;315(5814):961-6. PubMed PMID: 17303747.

12. Neunuebel JP, Knierim JJ. CA3 retrieves coherent representations from degraded input: direct evidence for CA3 pattern completion and dentate gyrus pattern separation. Neuron. 2014 Jan 22;81(2):416-27. PubMed PMID: 24462102. Pubmed Central PMCID: 3904133.

13. Deng W, Mayford M, Gage FH. Selection of distinct populations of dentate granule cells in response to inputs as a mechanism for pattern separation in mice. eLife. 2013;2:e00312. PubMed PMID: 23538967. Pubmed Central PMCID: PMC3602954. Epub 2013/03/30. eng.

14. Marrone DF, Adams AA, Satvat E. Increased pattern separation in the aged fascia dentata. Neurobiology of aging. 2011 Dec;32(12):2317 e23-32. PubMed PMID: 20447731. Epub 2010/05/08. eng.

15. Nakazawa K. Dentate Mossy Cell and Pattern Separation. Neuron. 2017;93(3):465-7.

16. Myers CE, Scharfman HE. A role for hilar cells in pattern separation in the dentate gyrus: a computational approach. Hippocampus. 2009 Apr;19(4):321-37. PubMed PMID: 18958849. Pubmed Central PMCID: 2723776.

17. Yim MY, Hanuschkin A, Wolfart J. Intrinsic rescaling of granule cells restores pattern separation ability of a dentate gyrus network model during epileptic hyperexcitability. Hippocampus. 2014 Oct 1. PubMed PMID: 25269417. Epub 2014/10/02. Eng.

18. Chavlis S, Petrantonakis PC, Poirazi P. Dendrites of dentate gyrus granule cells contribute to pattern separation by controlling sparsity. Hippocampus. 2017.

19. Neher T, Cheng S, Wiskott L. Memory storage fidelity in the hippocampal circuit: the role of subregions and input statistics. PLoS computational biology. 2015 May;11(5):e1004250. PubMed PMID: 25954996. Pubmed Central PMCID: PMC4425359. Epub 2015/05/09. eng.

20. Chawla MK, Guzowski JF, Ramirez-Amaya V, Lipa P, Hoffman KL, Marriott LK, et al. Sparse, environmentally selective expression of Arc RNA in the upper blade of the rodent fascia

650 dentata by brief spatial experience. Hippocampus. 2005;15(5):579-86. PubMed PMID:
651 15920719. Epub 2005/05/28. eng.
652 21. Ramirez S, Liu X, Lin PA, Suh J, Pignatelli M, Redondo RL, et al. Creating a false
653 memory in the hippocampus. Science. 2013 Jul 26;341(6144):387-91. PubMed PMID:
654 23888038.
655 22. Kelemen E, Fenton AA. Coordinating different representations in the hippocampus.
656 Neurobiology of learning and memory. 2016 Mar;129:50-9. PubMed PMID: 26748023. Epub
657 2016/01/10. eng.
658 23. Buzsaki G. Neural syntax: cell assemblies, synapsembles, and readers. Neuron. 2010 Nov
659 4;68(3):362-85. PubMed PMID: 21040841. Pubmed Central PMCID: 3005627.
660 24. Eichenbaum H. Time cells in the hippocampus: a new dimension for mapping memories.
661 Nature reviews Neuroscience. 2014 Nov;15(11):732-44. PubMed PMID: 25269553. Pubmed
662 Central PMCID: PMC4348090. Epub 2014/10/02. eng.
663 25. Rieke F. WW, de Ruyter van Steveninck R., Bialek W. Spikes: Exploring the Neural
664 Code. PRESS TM, editor: A Bradford Book; 1999.
665 26. VanRullen R, Guyonneau R, Thorpe SJ. Spike times make sense. Trends in
666 neurosciences. 2005 Jan;28(1):1-4. PubMed PMID: 15626490. Epub 2005/01/01. eng.
667 27. Gutig R. To spike, or when to spike? Current opinion in neurobiology. 2014 Apr;25:134-
668 9. PubMed PMID: 24468508. Epub 2014/01/29. eng.
669 28. Hopfield JJ. Neural networks and physical systems with emergent collective
670 computational abilities. Proceedings of the National Academy of Sciences of the United States of
671 America. 1982 Apr;79(8):2554-8. PubMed PMID: 6953413. Pubmed Central PMCID:
672 PMC346238. Epub 1982/04/01. eng.

673 29. Kreuz T, Chicharro D, Houghton C, Andrzejak RG, Mormann F. Monitoring spike train
674 synchrony. Journal of neurophysiology. 2013 Mar;109(5):1457-72. PubMed PMID: 23221419.
675 Epub 2012/12/12. eng.
676 30. Faisal AA, Selen LP, Wolpert DM. Noise in the nervous system. Nature reviews
677 Neuroscience. 2008 Apr;9(4):292-303. PubMed PMID: 18319728. Pubmed Central PMCID:
678 PMC2631351. Epub 2008/03/06. eng.
679 31. Dobrunz LE, Stevens CF. Response of hippocampal synapses to natural stimulation
680 patterns. Neuron. 1999 Jan;22(1):157-66. PubMed PMID: 10027298. Epub 1999/02/23. eng.
681 32. Aimone JB, Wiles J, Gage FH. Computational influence of adult neurogenesis on
682 memory encoding. Neuron. 2009 Jan 29;61(2):187-202. PubMed PMID: 19186162. Pubmed
683 Central PMCID: 2670434.
684 33. Savanthrapadian S, Meyer T, Elgueta C, Booker SA, Vida I, Bartos M. Synaptic
685 Properties of SOM- and CCK-Expressing Cells in Dentate Gyrus Interneuron Networks. The
686 Journal of neuroscience : the official journal of the Society for Neuroscience. 2014 Jun
687 11;34(24):8197-209. PubMed PMID: 24920624.
688 34. Ewell LA, Jones MV. Frequency-tuned distribution of inhibition in the dentate gyrus.
689 The Journal of neuroscience : the official journal of the Society for Neuroscience. 2010 Sep
690 22;30(38):12597-607. PubMed PMID: 20861366. Epub 2010/09/24. eng.
691 35. Hu H, Gan J, Jonas P. Interneurons. Fast-spiking, parvalbumin(+) GABAergic
692 interneurons: from cellular design to microcircuit function. Science. 2014 Aug
693 1;345(6196):1255263. PubMed PMID: 25082707.
694 36. Neunuebel JP, Knierim JJ. Spatial firing correlates of physiologically distinct cell types
695 of the rat dentate gyrus. The Journal of neuroscience : the official journal of the Society for

Neuroscience. 2012 Mar 14;32(11):3848-58. PubMed PMID: 22423105. Pubmed Central PMCID: 3321836.

37. Senzai Y, Buzsáki G. Physiological properties and behavioral correlates of hippocampal granule cells and mossy cells. Neuron. 2017;93(3):691-704. e5.

38. GoodSmith D, Chen X, Wang C, Kim SH, Song H, Burgalossi A, et al. Spatial representations of granule cells and mossy cells of the dentate gyrus. Neuron. 2017;93(3):677-90. e5.

39. Danielson NB, Turi GF, Ladow M, Chavlis S, Petrantonakis PC, Poirazi P, et al. In Vivo Imaging of Dentate Gyrus Mossy Cells in Behaving Mice. Neuron. 2017;93(3):552-9. e4.

40. Hyde RA, Strowbridge BW. Mnemonic representations of transient stimuli and temporal sequences in the rodent hippocampus in vitro. Nature neuroscience. 2012 Oct;15(10):1430-8. PubMed PMID: 22960934. Pubmed Central PMCID: 3614351.

41. Zylberberg J, Hyde RA, Strowbridge BW. Dynamics of robust pattern separability in the hippocampal dentate gyrus. Hippocampus. 2016 May;26(5):623-32. PubMed PMID: 26482936. Epub 2015/10/21. eng.

42. Chavlis S, Poirazi P. Pattern separation in the hippocampus through the eyes of computational modeling. Synapse (New York, NY). 2017 Mar 18. PubMed PMID: 28316111. Epub 2017/03/21. eng.

43. Cutts CS, Eglen SJ. Detecting pairwise correlations in spike trains: an objective comparison of methods and application to the study of retinal waves. The Journal of neuroscience : the official journal of the Society for Neuroscience. 2014 Oct 22;34(43):14288-303. PubMed PMID: 25339742. Pubmed Central PMCID: PMC4205553. Epub 2014/10/24. eng.

718 44. Lyttle D, Fellous JM. A new similarity measure for spike trains: sensitivity to bursts and
719 periods of inhibition. Journal of neuroscience methods. 2011 Aug 15;199(2):296-309. PubMed
720 PMID: 21600921. Pubmed Central PMCID: PMC4120777. Epub 2011/05/24. eng.
721 45. Kreuz T. Measures of spike train synchrony. Scholarpedia [Internet]. 2011; 6(10).
722 Available from: http://www.scholarpedia.org/article/Measures_of_spike_train_synchrony.
723 46. Pernía-Andrade Alejandro J, Jonas P. Theta-Gamma-Modulated Synaptic Currents in
724 Hippocampal Granule Cells In Vivo Define a Mechanism for Network Oscillations. Neuron.
725 2014;81(1):140-52.
726 47. Diamantaki M, Frey M, Berens P, Preston-Ferrer P, Burgalossi A. Sparse activity of
727 identified dentate granule cells during spatial exploration. eLife. 2016 Oct 03;5. PubMed PMID:
728 27692065. Epub 2016/10/04. Eng.
729 48. Lisman JE, Jensen O. The theta-gamma neural code. Neuron. 2013 Mar 20;77(6):1002-
730 16. PubMed PMID: 23522038. Pubmed Central PMCID: 3648857.
731 49. Wikenheiser AM, Redish AD. Decoding the cognitive map: ensemble hippocampal
732 sequences and decision making. Current opinion in neurobiology. 2015;32:8-15.
733 50. Larimer P, Strowbridge BW. Representing information in cell assemblies: persistent
734 activity mediated by semilunar granule cells. Nature neuroscience. 2010 Feb;13(2):213-22.
735 PubMed PMID: 20037579. Pubmed Central PMCID: 2840722.
736 51. Friedrich RW, Wiechert MT. Neuronal circuits and computations: pattern decorrelation
737 in the olfactory bulb. FEBS letters. 2014 Aug 1;588(15):2504-13. PubMed PMID: 24911205.
738 Epub 2014/06/10. eng.
739 52. Fontanari J, Rolls E, Costa DF. A model of the operation of the hippocampus and
740 entorhinal cortex in memory. International Journal of Neural Systems. 1995;6(1):51-70.

741 53. Aimone JB, Deng W, Gage FH. Resolving new memories: a critical look at the dentate
742 gyrus, adult neurogenesis, and pattern separation. Neuron. 2011 May 26;70(4):589-96. PubMed
743 PMID: 21609818. Pubmed Central PMCID: PMC3240575. Epub 2011/05/26. eng.
744 54. Wick SD, Wiechert MT, Friedrich RW, Riecke H. Pattern orthogonalization via channel
745 decorrelation by adaptive networks. Journal of computational neuroscience. 2010 Feb;28(1):29-
746 45. PubMed PMID: 19714457. Epub 2009/08/29. eng.
747 55. Teyler TJ, Rudy JW. The hippocampal indexing theory and episodic memory: updating
748 the index. Hippocampus. 2007;17(12):1158-69. PubMed PMID: 17696170. Epub 2007/08/19.
749 eng.
750 56. Severa W, Parekh O, James CD, Aimone JB. A Combinatorial Model for Dentate Gyrus
751 Sparse Coding. Neural computation. 2016 Oct 20:1-24. PubMed PMID: 27764589. Epub
752 2016/10/21. Eng.
753 57. Bischofberger J, Engel D, Li L, Geiger JR, Jonas P. Patch-clamp recording from mossy
754 fiber terminals in hippocampal slices. Nature protocols. 2006;1(4):2075-81. PubMed PMID:
755 17487197. Epub 2007/05/10. eng.
756 58. Harney SC, Jones MV. Pre- and postsynaptic properties of somatic and dendritic
757 inhibition in dentate gyrus. Neuropharmacology. 2002 Sep;43(4):584-94. PubMed PMID:
758 12367604. Epub 2002/10/09. eng.
759 59. Andersen P, Holmqvist B, Voorhoeve PE. Excitatory synapses on hippocampal apical
760 dendrites activated by entorhinal stimulation. Acta physiologica Scandinavica. 1966
761 Apr;66(4):461-72. PubMed PMID: 5927272. Epub 1966/04/01. eng.

762 60. Macke JH, Berens P, Ecker AS, Tolias AS, Bethge M. Generating spike trains with
763 specified correlation coefficients. Neural computation. 2009 Feb;21(2):397-423. PubMed PMID:
764 19196233. Epub 2009/02/07. eng.
765 61. Danielson NB, Kaifosh P, Zaremba JD, Lovett-Barron M, Tsai J, Denny CA, et al.
766 Distinct contribution of adult-born hippocampal granule cells to context encoding. Neuron.
767 2016;90(1):101-12.
768 62. Dieni CV, Panichi R, Aimone JB, Kuo CT, Wadiche JI, Overstreet-Wadiche L. Low
769 excitatory innervation balances high intrinsic excitability of immature dentate neurons. Nature
770 communications. 2016;7:11313. PubMed PMID: 27095423. Pubmed Central PMCID:
771 PMC4843000. Epub 2016/04/21. eng.
772 63. Marin-Burgin A, Mongiat LA, Pardi MB, Schinder AF. Unique processing during a
773 period of high excitation/inhibition balance in adult-born neurons. Science. 2012 Mar
774 9;335(6073):1238-42. PubMed PMID: 22282476. Pubmed Central PMCID: PMC3385415. Epub
775 2012/01/28. eng.
776 64. de la Rocha J, Doiron B, Shea-Brown E, Josic K, Reyes A. Correlation between neural
777 spike trains increases with firing rate. Nature. 2007 Aug 16;448(7155):802-6. PubMed PMID:
778 17700699. Epub 2007/08/19. eng.
779 65. Aronov D, Victor JD. Non-Euclidean properties of spike train metric spaces. Physical
780 Review E. 2004 06/02/;69(6):061905.
781 66. Motulsky HJ, Ransnas LA. Fitting curves to data using nonlinear regression: a practical
782 and nonmathematical review. FASEB journal : official publication of the Federation of
783 American Societies for Experimental Biology. 1987 Nov;1(5):365-74. PubMed PMID: 3315805.
784 Epub 1987/11/01. eng.

785

Fig 1. Pattern separation assay in acute brain slices at the single cell level.

(A) Five input sets out of eleven used. *Top*: rasters of the five spiketrains of each set. *Bottom*: correlation coefficient matrix for each input set, each square representing the correlation coefficient between two input spiketrains measured with a binning window of 10 ms (color scale at left). Traces are ordered by decreasing similarity (i.e. coefficient average, diagonal excluded) from R_{input} = 0.95 (far left) to R_{input}=0.11 (far right).

(B) Histology of the DG in a horizontal slice (Cresyl violet/Nissl staining; scale bar: 250µm), overlaid with a schematic of the experimental setup: a theta pipette in the ML (input) is used to focally stimulate the PP while a responding GC (output) is recorded via whole-cell patch-clamp. (GCL: granule cell layer, H: hilus, ML: molecular layer, FS: fast-spiking interneuron. Solid lines represent dendrites and dashed lines axons)

(C) Current-clamp recordings of the membrane potential of two different GCs in response to different input sets (*Top*: R_{input} = 1; *Bottom*: R_{input} = 0.76). Each set of five input traces is repeated ten times (only 3 repetitions are shown, with spikes truncated at 0 mV). In the bottom graph, input trains and their respective children output spiketrains have matching colors.

Fig 2. Input spiketrains are decorrelated at the level of individual granule cells.

(A) Example of a recording set (input set + output set): the raster plot shows one set of input spiketrains and the children output spiketrains recorded from one GC, organized so that output subsets (i.e., the ten children coming from one parent input spiketrain) are together and of the same color.

(B) Corresponding 55x55 correlation coefficient matrix using a binning window (τ_w) of 10 ms. Each small square represents the correlation coefficient between two spiketrains. R_{output} is defined as the mean of correlations between individual output spiketrains driven by different input spiketrains, as outlined by the bold blue border, which excludes comparisons between outputs generated from the same parent input.

(C) For each of the 102 recording sets (blue dots), R_{output} was lower than $R_{input.}$ Black dots and error bars represent means and SEM (as in D and E).

(D) *Left*: effective decorrelation averaged over all recording sets as a function of R_{input}. Although there is a significant decorrelation for all tested input sets (one-sample T-tests: the blue shade indicates the 95% confidence interval that average decorrelation is significantly above 0), they are effectively decorrelated to different magnitudes (one-way ANOVA, $p<0.0001$). *Right:* post-hoc Tukey-Kramer tests show that the decorrelation is significantly different (higher) for highly similar input spiketrains than for already dissimilar inputs.

(E) When the effective decorrelation is normalized to the correlation of the input set, there is no significant difference between input sets (ANOVA, $p = 0.18$).

In all graphs, $\tau_w = 10$ ms.

Fig 3. Orthogonalization of input spiketrains is a strong component of temporal pattern separation by single granule cells.

(A-C) Three hypothetical cases of pairs of spiketrains and their associated correlation coefficients (R), normalized dot products (NDP) and scaling factors (SF), showing that the three metrics are not equivalent.

(A) Synthetic spiketrains (X and Y pairs) divided into six bins, with the corresponding number of spikes per bin.

(B) R between each pair of X and Y describes the linear regression between the number of spikes in the bins of X versus the corresponding bins in Y (jitter was added to make all points visible).

(C) Geometric view of vectors X and Y, where each bin is a dimension of a 6-dimensional space, and the number of spikes in a bin is the coordinate along this dimension. NDP measures how close to orthogonal two spiketrains are and SF measures how different their binwise firing rates are.

(D) Vector representation of experimental data from one recording set, showing the average similarity between a set of input spiketrains (dashed line and green angle, $R_{input} = 0.76$) and the average similarity between the fifty corresponding output spiketrains (solid line, purple angle, over the same subsets as enclosed in the blue border in **Fig 2B**). The angles are derived from the NDP whereas the lengths of each vector express differences in binwise firing rates (SF). Here, outputs are more orthogonal (closer to 90˚) than their inputs with little difference in scaling.

(E-F) NDP_{output} or SF_{output} as a function of NDP_{input} or SF_{input}. Mean and SEM in black.

(E) All data points (102 recording sets) are below the dashed identity line indicating that outputs are closer to orthogonality (NDP = 0) than their respective inputs. The average orthogonalization (NDP_{input}-NDP_{output}) is significant for all input sets (one-sample T-tests, $p < 0.05$).

(F) Most points fall slightly below the dashed identity line, suggesting that pattern separation by scaling of the binwise firing rate is present but weak at τ_w = 10ms. The average scaling (SF_{input}-SF_{output}) is significant for all input sets except the three most dissimilar (SF_{input} = 0.88, 0.89, 0.90) (one-sample T-tests, $p < 0.05$).

Fig 4. Single granule cells perform pattern separation on millisecond to second timescales using different codes.

(A) *Top:* Average R_{output} as a function of R_{input}, measured with different time windows τ_w. Solid curves are fitted parabolae. Each color corresponds to a different τ_w ranging from 5 ms to 100 ms. B*ottom:* Effect of τ_w on the effective decorrelation, interpolated from the parabolic regressions.

(B) Same as A but using NDP and linear regressions. *Bottom:* Note that the timescale axis is extended to 2000 ms, with the inset showing an expansion over the shorter timescales).

(C) Same as B using SF.

(B-C) Note that as τ_w increases, pattern separation through orthogonalization becomes weaker but stronger through scaling.

(D) Similarity between spiketrains is here assessed with the binless SPIKE metric, directly using spike times. *Left:* example of two input spiketrains associated with two output spiketrains from a GC recording set, and the corresponding distances D(t) between spiketrains. D(t) can then be integrated over time to give a single value D. *Middle*: example of 55x55 matrix of SPIKE similarity (1-D) between all spiketrains of an example recording set. 0 means that spikes of two trains never happened close in time, and 1 that they were perfectly synchronous. The output SPIKE similarity ($SPIKE_{output}$) is defined the same way as for R, NDP or SF (average of the values inside the blue border). *Right*: $SPIKE_{output}$ of all GC recordings as a function of their input similarity ($SPIKE_{input}$), fitted with a parabola (red line). Most data points are below the dashed identity line indicating that output spiketrains are less similar than inputs. The average $SPIKE_{input}$-$SPIKE_{output}$ is significantly above 0 for all input sets except the two most dissimilar ($SPIKE_{input}$ = 0.74, 0.78) (one-sample T-tests, $p < 0.05$).

Fig 5. Input spiketrains are efficiently separated upon their first presentations.

(A) Two of five inputs are shown with corresponding output spiketrains. The first output sweep is marked with a pink bar (right) and last sweep is marked with a blue bar.

(B) R_{output}, computed from the first sweep of five output trains only (pink), as a function of R_{input}, fitted with a parabola. All data points are below the identity line indicating that outputs are effectively decorrelated compared to their inputs even when input patterns have only been presented once each. The average decorrelation (R_{input}-R_{output}) is significant for all input sets (one-sample T-tests, $p < 0.01$) except for $R_{input} = 0.11$ ($p = 0.1$).

(C) *Left*: Average output correlations between spiketrains of the first sweep (pink) and the last sweep (blue). There is no significant difference (F-test comparing the two distributions using parabolic regressions, $p = 0.47$. Error bars are SEM). *Right:* When taking into account that the two distributions are paired, we detect that a few output correlations are significantly lower for the last sweep than for the first one (one-sample T-test on the difference between R_{output} of the first and last sweep of each recording set, asterisks signify $p < 0.05$). This is evidence, though weak, that repetition of input spiketrains might improve pattern separation for highly similar inputs.

Fig 6. Pattern separation in single GCs is not explained by simple neural noise.

(A) The variability of output spiketrains in response to the same input train sets the upper bound for R_{output}. *Left:* Correlations between pairs of output spiketrains associated with different input trains (enclosed by red, R_{output}) and pairs of different output spiketrains associated with the same input train (enclosed by green, R_w: spiketrain reliability, the reproducibility of the output given the same input). *Right:* Empirical probability distribution of R_w for all recordings (dark green line is the mean: $\langle R_w \rangle = 0.3$), overlaid on the distribution of R_{output} as a function of R_{input} (102 recording sets) fitted with a parabola. Note that means $\langle R_w \rangle$ and $\langle R_{output} \rangle$ for $R_{input} = 1$ are close because they both assess the reproducibility of the output when the input is the same.

(B) Characterization of neural noise. *Top:* example of input and output spiketrains illustrating variable delay of the response spike after an input spike (d1<d2) or failure to spike after an input spike (red cross). *Bottom*: Example from one GC recording. The spike-wise noise in output spiketrains is characterized by the average spike delay, the standard deviation of this delay (jitter) and the probability of spiking after an input spike (spiking reliability of the cell, SR).

(C-D) Effect of random shuffling on R_{output} and R_w.

(C) Patterns are *less* separated for GC than for random shuffling (F-test using parabolic regressions: $p < 0.0001$). (D) GC output is *more* reliable than for random shuffling (unpaired T-test, $p < 0.0001$; mean $\langle Rw \rangle \{shuffle\} = 0.19$).

(E) Paired statistical tests show that shuffling leads to smaller R_{output} and R_w than original recordings. *Top:* paired T-test on all recording sets, $p < 0.0001$. *Bottom*: one-sample T-test on difference between shuffled and original data R_{output} performed on each R_{input} group. Black symbols correspond to the means, bars to SEM. The purple shade indicates the 95% confidence interval. Asterisks signify $p < 0.001$.

Fig 7. Unreliability in spiketrain transmission is a major but not unique source of temporal pattern separation.

(A) Spiketrain reliability (R_W) is an excellent predictor of normalized decorrelation (defined in **Figure 2E**). Green: 102 recording sets from GC recording sets; Red: 20 recording sets from fast-spiking interneurons (FS). Notice that, despite the strong anti-correlation, the affine model predicts that even a perfect reliability ($R_W = 1$) could still allow 10% of decorrelation. See **Table 3** for linear regressions on GCs or FS alone.

(B) The ten children output spiketrains of each of the five inputs can be averaged to give the five output peristimulus histograms (PSTH) corresponding to the five input trains. The 10 ms binned PSTHs of the output rasters in **Fig 2A** are shown.

(C) Correlation coefficients between all pairs of the five output PSTHs. The mean correlation (PSTH R_{output}) is the average of coefficients inside the red border, and excludes self-comparisons.

(D) *Left:* PSTH R_{output} as a function of R_{input} (102 recording sets, in red), fitted with a parabola (black). All points are below the identity line indicating decorrelation of outputs compared to inputs. *Right:* Average effective decorrelation (R_{input} – PSTH R_{output}) as a function of R_{input} (bars are SEM) reveals a significant decorrelation for all input sets except for the most dissimilar (one-sample T-tests; shaded area is the 95% confidence interval for significant decorrelation).

(E) Averaged PSTH R_{output} as a function of R_{input}, for different binning windows τ_w.

Fig 8. Fast-spiking interneurons of the DG exhibit lower levels of pattern separation than GCs.

(A) Picture of a recorded FS filled with biocytin (black). In the case of simultaneous recordings, the recorded GCs were close to the FS, as depicted by the schematic in green.

(B-C) Example of a simultaneous whole-cell recording of a GC and a neighboring FS.

(B) Simultaneous membrane potential recordings (baseline around -60mV) of a FS and a GC to the same set of current steps (-25pA, 100pA, 500pA and 1000pA).

(C) Simultaneous current-clamp recordings of the same FS and GC as in A in response to the five input traces of an input set with $R_{input} = 0.65$ (first sweep of five output trains). Simultaneous input and output trains have the same color.

(D) R_{output} versus R_{input} at $\tau_w = 10$ms. Data points correspond to recording sets: 20 for FS (red), and 61 for GC (green, with a darker shade open circle when simultaneously recorded with a FS). All GC recordings done at the same input correlations as FS recordings were used for an unpaired comparison: R_{output} distribution for FS is significantly higher in FS than in GC (F-test using linear regressions: $p < 0.0001$ (asterisk); unbalanced two-way ANOVA, stimulation groups: $p = 0.0015$, cell-types: $p < 0.0001$, interaction: $p = 0.72$.) Post-hoc Tukey-Kramer tests: significant difference for stimulation groups with measured correlation $R_{input} = 0.88$, 0.84 and 0.74 ($p < 0.01$).

(E) Spiketrain reliability R_W. Same color code and recording sets as in D**.** Unpaired T-test: $p < 0.0001$ (asterisk).

(**D-E**) Note that when comparing only the simultaneous GC and FS recordings, we found a similarly significant difference, both for R_{output} and R_w.

Fig 9. Cell-to-cell comparisons show that granule cells have variable responses to identical inputs and perform pattern separation in different ways.

(A) Spiketrain-wise comparison between recording sets.

(A1) The similarity between pairs of spiketrains coming from two different output sets but associated to the same input set and with the same sweep number is assessed with the Pearson's correlation coefficient (τ_w = 10 ms). The fifty resulting coefficients are then averaged to give $R_{cell\text{-}to\text{-}cell}$, a single number measuring the overall similarity of output spiketrains between two recording sets. All combinations of pairs of output sets from the same input set were compared.

(A2) Probability distribution of $R_{cell\text{-}to\text{-}cell}$ (green line) across all GC recordings. The distribution of $R_{cell\text{-}to\text{-}cell}$ (black circles) is not dependent on R_{input}.

(A3) Distribution of $R_{cell\text{-}to\text{-}cell}$ for different cell-types. (GC to GC: n = 470 , FS to FS: n = 30, FS to GC simultaneously recorded: n = 15).

(B) Comparison of the pattern separation levels between recording sets.

(B1) Top: same as Fig. 2B but in grey scale. Colored squares enclose the coefficients comparing outputs in response to two different input trains which have a $R_{input} \sim 0.76$ shown in (A1). Bottom: Average decorrelation of input spiketrains for each group of comparisons enclosed by the matching colored square above. Here, for each of these groups, R_{output} is the average of the coefficients in the square of the corresponding color and R_{input} is the correlation coefficient between the corresponding two parent input trains (not the average R over all input trains). For each recording set, a one-way ANOVA was performed to compare the 10 groups. For the GC shown, the ANOVA was significant, suggesting that levels of decorrelation depend on the identity of the input trains.

(B2) Percentage of recording sets with a significant ANOVA (crosses in black, axis on left). For all input sets except $R_{input} = 1$ the proportion was high. The distribution of the largest difference between the mean of the ten decorrelation groups of a single recording set is also plotted (grey circles, axis on right).

(B3) To assess whether cells perform similar levels of pattern separation we computed $R_{cell\text{-}to\text{-}cell}\{decorr\}$, the Pearson's correlation coefficient between decorrelation groups (as in B1) from pairs of recording sets.

(B4) Distribution of $R_{cell\text{-}to\text{-}cell}\{decorr\}$ (GCs only, mean and SEM in red). Like in **A**, all combinations of pairs from the same input set were compared (442 total comparisons). $R_{input} = 1$ was excluded. This analysis shows that, in our experimental conditions, not all GC decorrelate input spiketrains the same way.

S1 Fig. The correlation coefficient, normalized dot product and scaling factor between spiketrains do not have a simple relationship.

(A-C) Two additional examples of hypothetical pairs (X and Y) of spiketrains divided into 6 bins, the number of spikes per bin constituting the numerical vector next to them (A right). The similarity between X and Y can be assessed by (B) the Pearson's correlation coefficient R describing the linear relationship between the number of spikes in corresponding bins or (C) by a vector analysis in a 6-dimensional space giving the angle between the vectors X and Y (computed from the normalized dot product, NDP) as well as the scaling factor (SF) between their norms. These examples provide the intuition that orthogonal vectors (NDP = 0) necessarily correspond to a negative correlation between the spiketrains but that anticorrelated spiketrains (R<0) are not necessarily orthogonal. Also, SF gives information about the firing rate per bin that R doesn't necessarily take into account (see second example in **Fig 3**).

(D) Relationships between these 3 similarity metrics (R, NDP and SF) computed between 1,000 randomly generated spiketrains with six bins like in (A) (All pairs combinations = 499,500 data points). For simplicity, each bin could have only 0, 1 or 2 spikes, which is why points in the SF graph are less distributed.

(E) Relationships between these three similarity metrics for 102 experimental GC recording sets. (124,950 data points)

(F) Relationship between these three similarity metrics averaged for each GC recording set (102 data points).

(E-F) Green lines correspond to a linear regression, the R^2 and the p-value of which are indicated in each panel. Note that although R and NDP are well correlated in our experimental data

(R^2>0.95) (E-F left), there is not a linear relationship between R and NDP in theory **(**D left and see **Materials and methods**).

S2 Fig. Spike delay, jitter and reliability distributions for real data, simulations and shuffled data

(A) Cross-occurrence method to measure spike delay, jitter and spiking reliability of a neuron during a given recording session. *Top*: Example histogram of output spikes occurring after input spikes, fitted (red curve) with a Gaussian distribution N(μ,σ, baseline), where μ is the mean delay and σ is the jitter of this delay. Lag 0 ms corresponds to the input spike time. In this example, output spikes are generated on average 16 ms after a stimulation impulse (delay) with a jitter (σ) of 8.7 ms. *Bottom***:** the baseline is subtracted and the histogram divided by the number of input spikes during the recording session. This gives the distribution of the probability of spiking after an input spike, the sum of probabilities defining the spiking reliability of the cell during the recording session. Here the neuron fires 39 % of the time after an input spike.

(B-D) Delay, jitter and spiking reliability (SR) distributions for (B) the original recordings, (C) the simulations and (D) the shuffled data set. Dashed horizontal red lines are means.

S3 Fig. Spike-wise neural noise characteristics are not good predictors of spiketrain decorrelation by single GCs.

Plots of the normalized decorrelation, i.e. $(R_{input} - R_{output}) / R_{input}$, of each recording set (τ_w = 10 ms) : 102 for GC original and shuffled recordings **(A-B)**, 20 for FS **(C)**, as a function of spike-wise noise characteristics (spike delay, jitter and reliability). Solid green lines are the best linear fit, with R^2 and p-values noted in each panel. These plots are examples to illustrate **Table 2**. Note that decorrelation is poorly explained (low R^2) by either the spike delay or its jitter in all cell-types. In contrast, the spiking reliability (SR) is a good predictor of decorrelation in shuffled GC recordings (i.e. recordings entirely dominated by spike-wise noise) and even more so in FS recordings (for FS, SR was computed from nbFS data). This suggests that SR can be a potent mechanism for decorrelation, and that FS show different levels of decorrelation than GCs because they are more reliable. However, SR is only a mediocre predictor of decorrelation for GCs, thus confirming that temporal pattern separation in single GCs cannot be the result of simple neural noise.

S4 Fig. Simulation of purely noisy outputs significantly decreases R_{output} and spiketrain reliability

Simulations of output spiketrains to the different input sets with random spiking following a Gaussian distribution defined by the mean spike delay, mean jitter and mean reliability of the original recordings.

(A) Pattern separation for different time windows τ_w (dots: average R_{output} across all simulated "recordings" for a given input set, bars are SEM).

(B) R_{output} distribution at $\tau_W = 10$ ms, for simulated data and original data. Distributions are fitted with parabolae and significantly different (F-test using parabolic regressions: $p < 0.0001$).

(C) R_w distributions are significantly different (unpaired T-test, $p < 0.0001$; <Rw>{simul} = 0.14).

(D) Like in the original data (**Fig 2E**), the average normalized decorrelation $((R_{input}-R_{output})/R_{input})$ seems invariant. Bars are SEM.

(C-D) Asterisks signify statistical significance

S5 Fig. Differences in pattern separation between FS and GC are not solely due to FS bursting behavior or higher firing rate.

(A) Probability of having 0, 1 or more output spikes between two input spikes assesses the bursting behavior in FS recordings.

(B) In contrast to GCs, for FS neurons there is a strong correlation between the firing rate of a recording set and the associated normalized decorrelation. See **Table 3** for linear regression goodness-of-fit and significance when considering GC only, or GC and FS combined.

(C) Example of bursts in a FS (*Bottom*) in response to input spikes (*Top*). To assess the effect of bursting on R_{output}, we truncated each recorded spiketrain from FS neurons to keep only the first output spike between two input spikes, thus removing any burst without altering the SR of the cell. The blue shaded areas highlight the spikes that were removed. The resulting truncated dataset was termed "nbFS" for "non-burst FS".

(D) The firing rate of nbFS neurons is significantly reduced compared to the firing rate of the original FS. (Kruskal-Wallis one-way ANOVA, $p < 0.0001$. FS vs nbFS (paired data): post-hoc Wilcoxon signed rank test with Bonferroni correction, $p < 0.001$; nbFS vs GC (non-paired): post-hoc Mann-Whitney U-test with Dunn-Sidak correction, $p = 0.02$; FS vs GC (non-paired): post-hoc Mann-Whitney U-test with Dunn-Sidak correction, $p < 0.001$).

(E) Distributions of R_{output} (τ_w = 10ms). Data points correspond to individual recording sets of nbFS (purple) or GC (green). Both distributions are still significantly different, suggesting the bursting behavior of FS is not sufficient to explain the difference in pattern separation (Unbalanced two-way ANOVA. Stimulation groups: $p = 0.002$, cell-types: $p < 0.0001$, interaction: $p = 0.84$. Post-hoc Tukey-Kramer tests: nbFS and GC R_{output} are significantly different ($p < 0.05$) for stimulation groups with measured correlation R_{input} = 0.88, 0.84, 0.74.)

(F) Distributions of R_w (τ_w = 10 ms) are still significantly different between GC and nbFS (unpaired T-test, $p < 0.0001$). This suggests the bursting behavior of FS is not sufficient to explain the difference in spiketrain reliability R_w.

(D-F) Asterisks signify $p < 0.001$.

S6 Fig. Fast-spiking interneurons and granule cells have different noise characteristics.

(A) Cross-occurrence method (same as in **S2 Fig**) to measure spike delay, jitter and spiking reliability of nbFS neurons (to only consider the noise characteristics of the first spike, if there is a burst) *Top*: example of a histogram representing the number of output spikes occurring after input spikes. The histogram is fitted (red curve) with a Gaussian distribution N(μ,σ, baseline), where μ is the mean delay and σ is the jitter of this delay. Lag 0 ms corresponds to the input spike time. In this example, output spikes are generated on average 7.8 ms after an input spike (delay) with a jitter of 4.2 ms. *Bottom***:** the baseline is subtracted and the histogram divided by the number of input spikes during the recording session. This gives the distribution of the probability of spiking after an input spike of the cell during the recording session (SR). Here the neuron fires 59% of the time after an input spike.

(B) Delay, jitter, and SR distributions for the 20 nbFS recording sets. Dashed horizontal black lines represent means. To compare with **S2 Fig**.

(C) Comparison of the delay, jitter and SR between nbFS and GC recordings (Mann-Whitney U-tests, $p < 0.0001$). FS first spike responses to a stimulation impulse is faster, has less jitter, and is more reliable than in GCs.

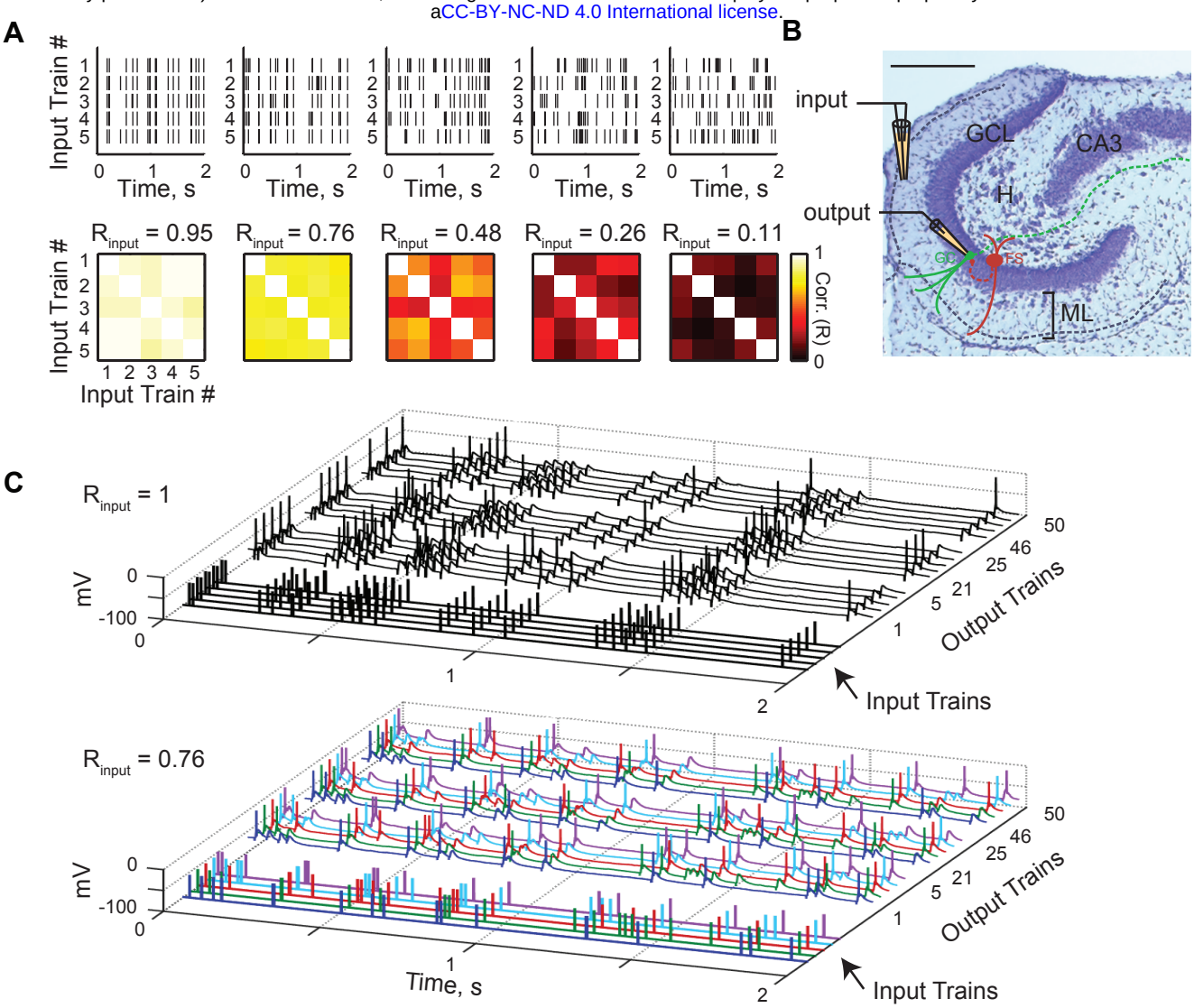


Figure 1

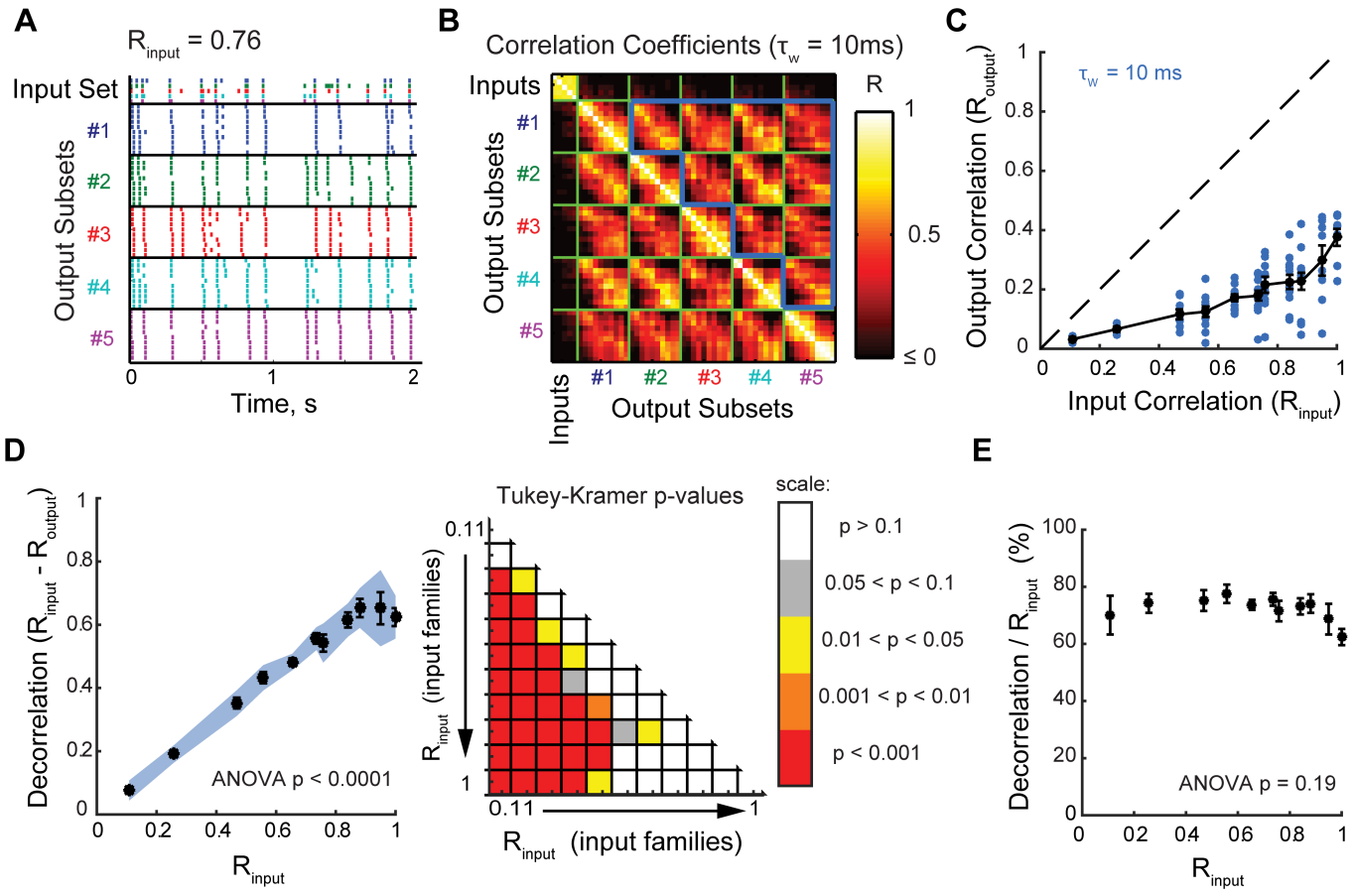


Figure 2

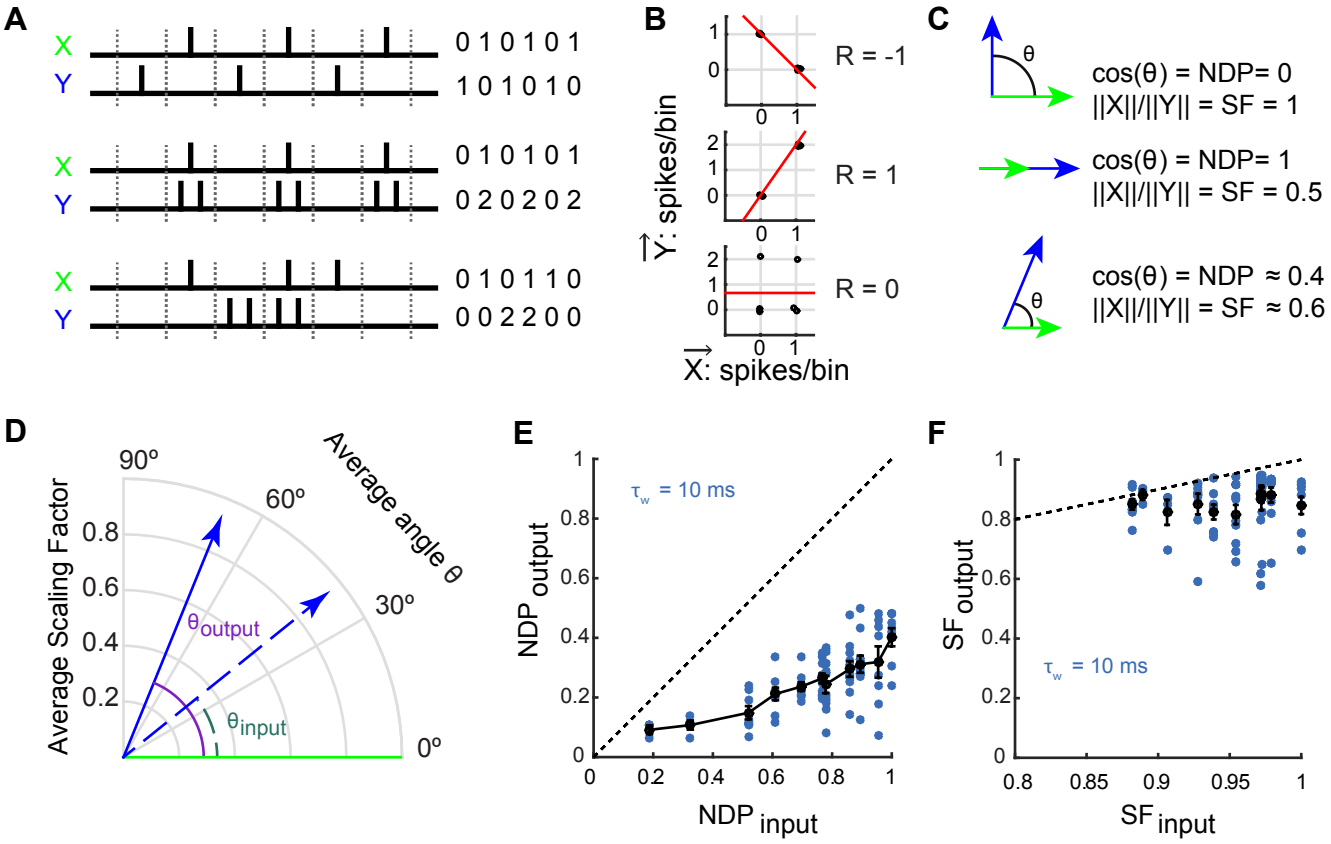


Figure 3

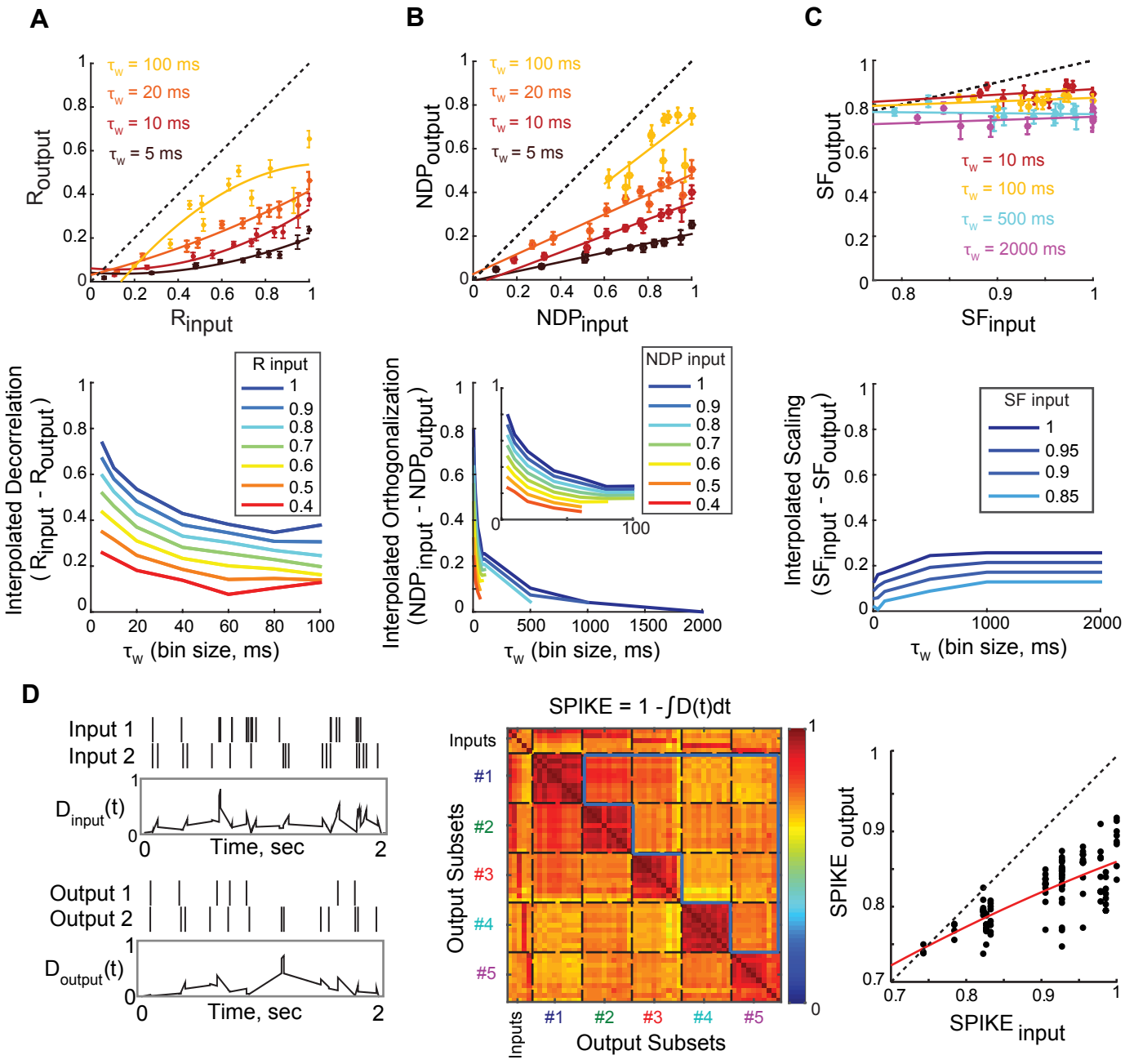


Figure 4

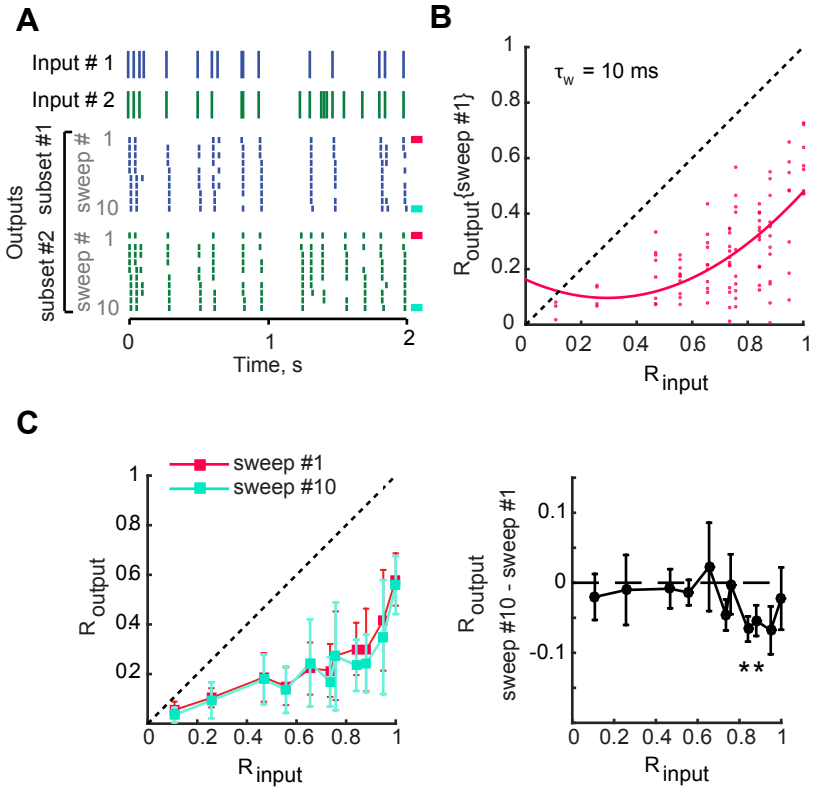


Figure 5

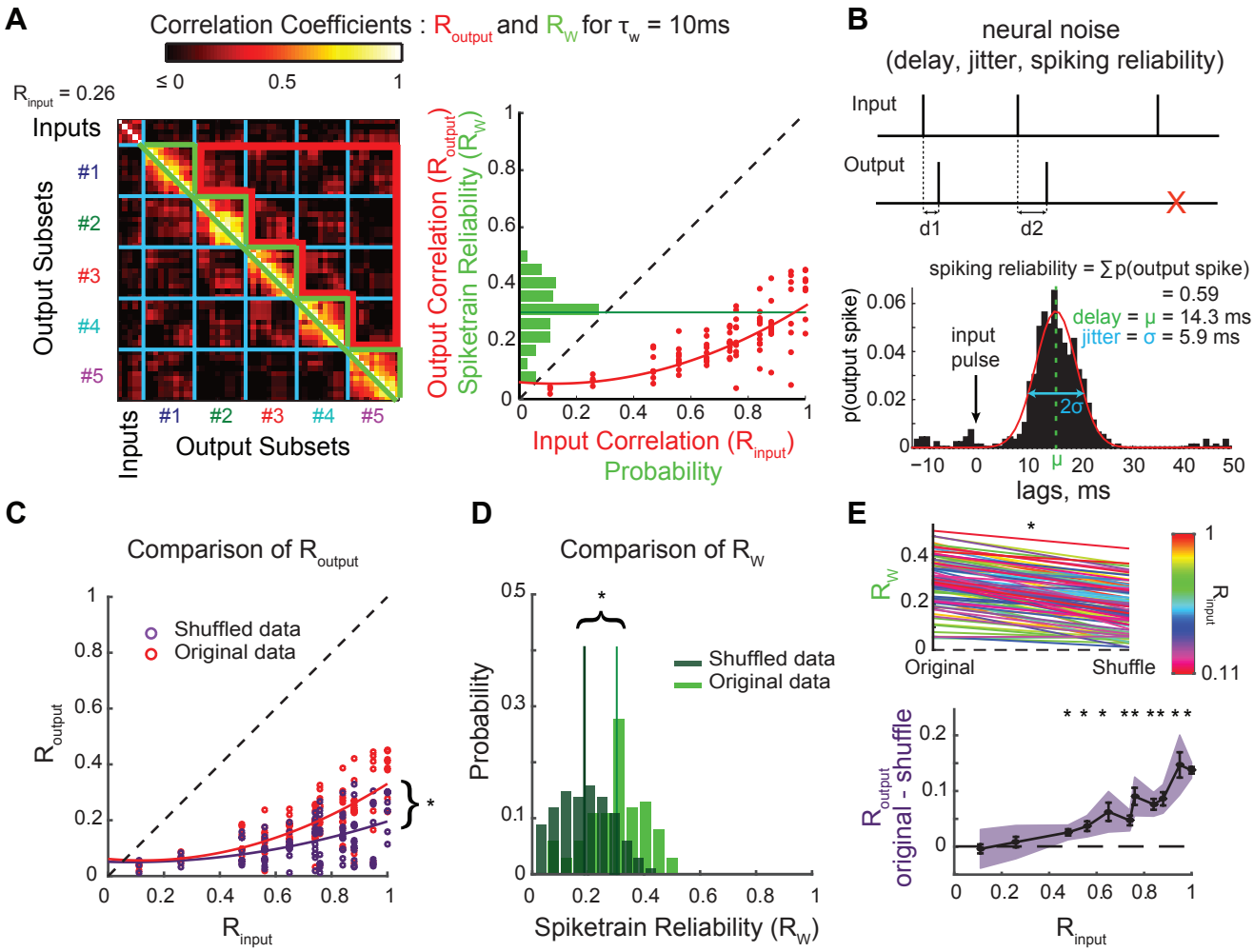


Figure 6

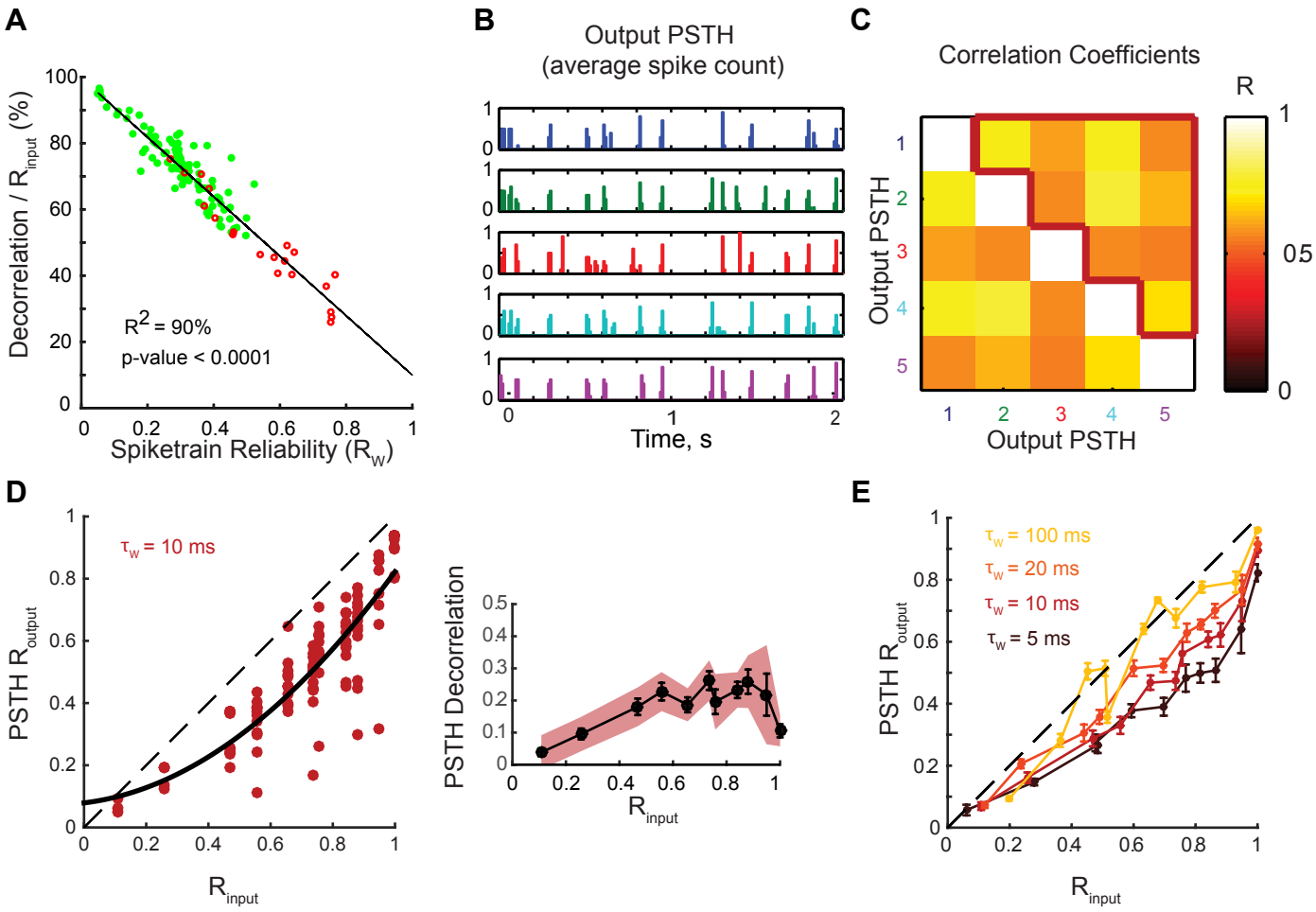


Figure 7

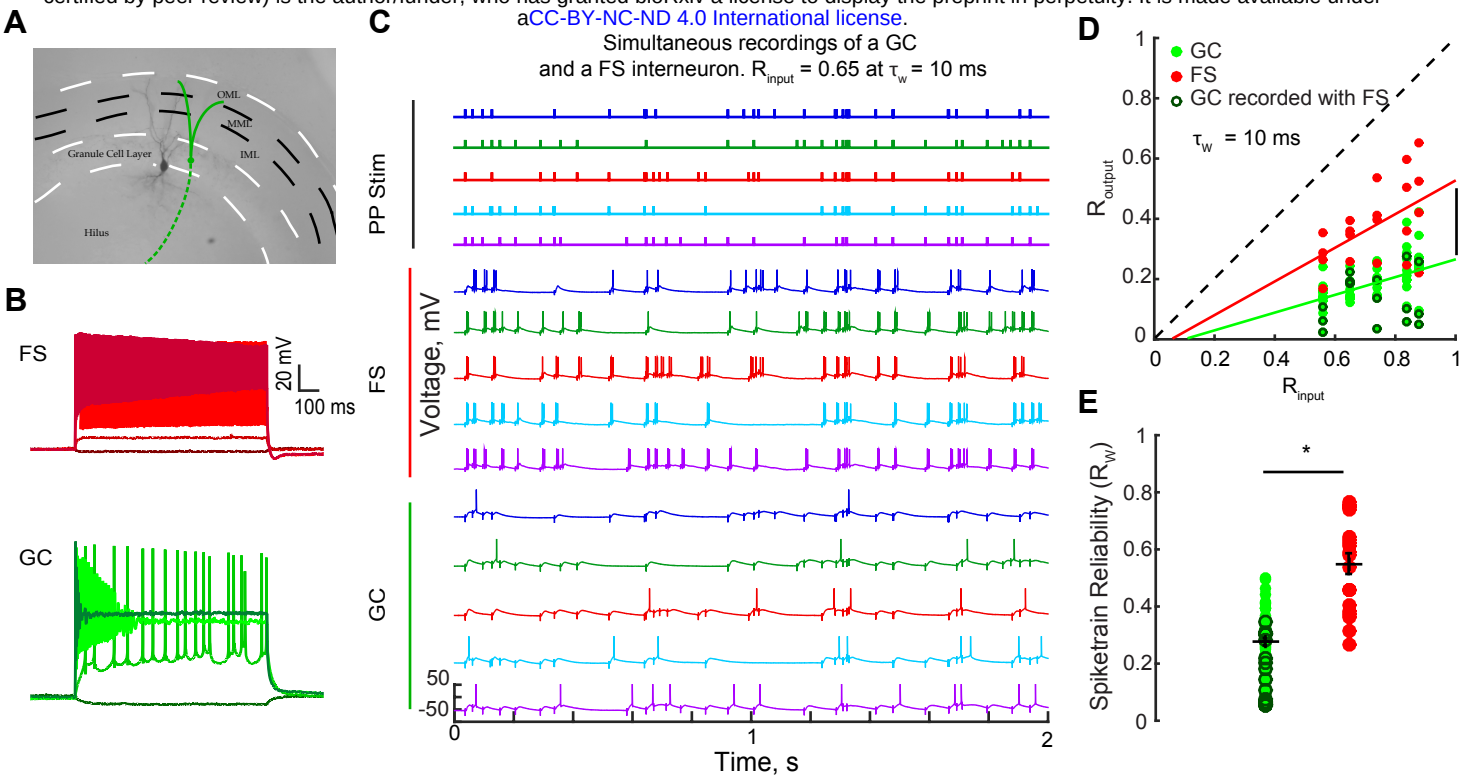


Figure 8

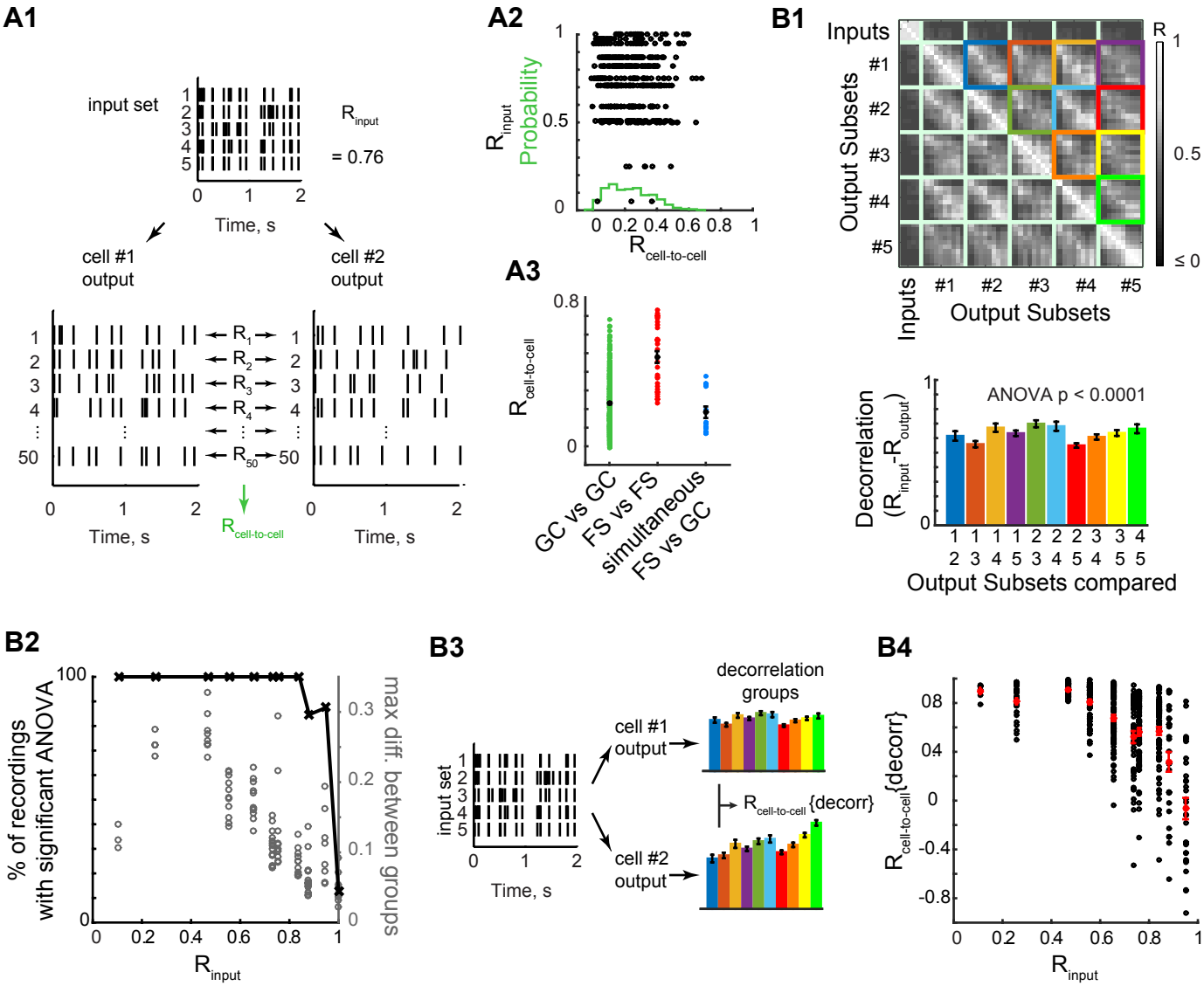
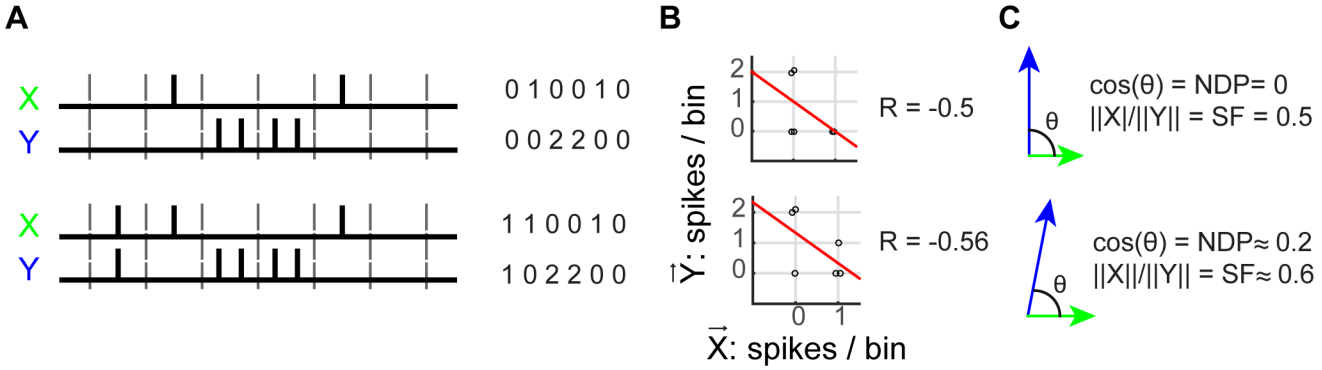
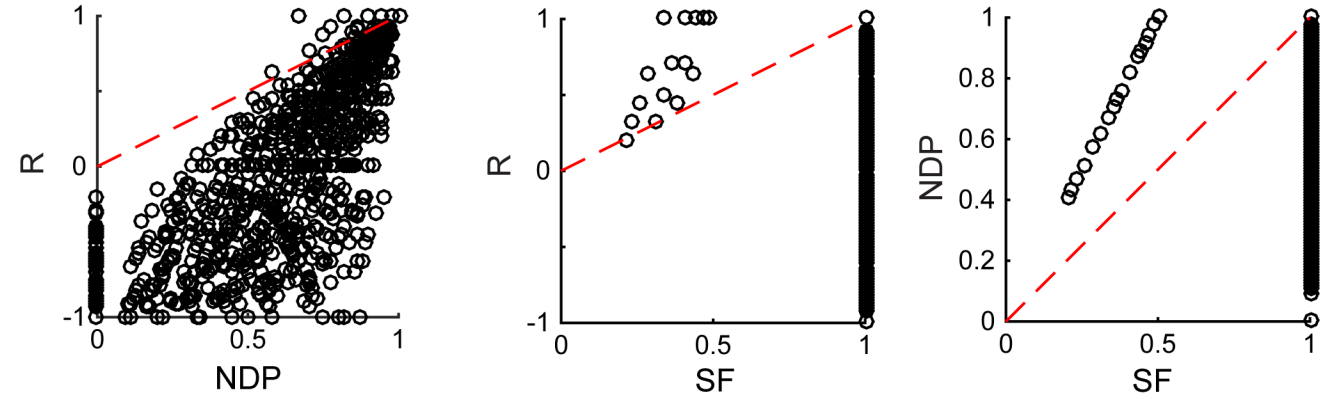


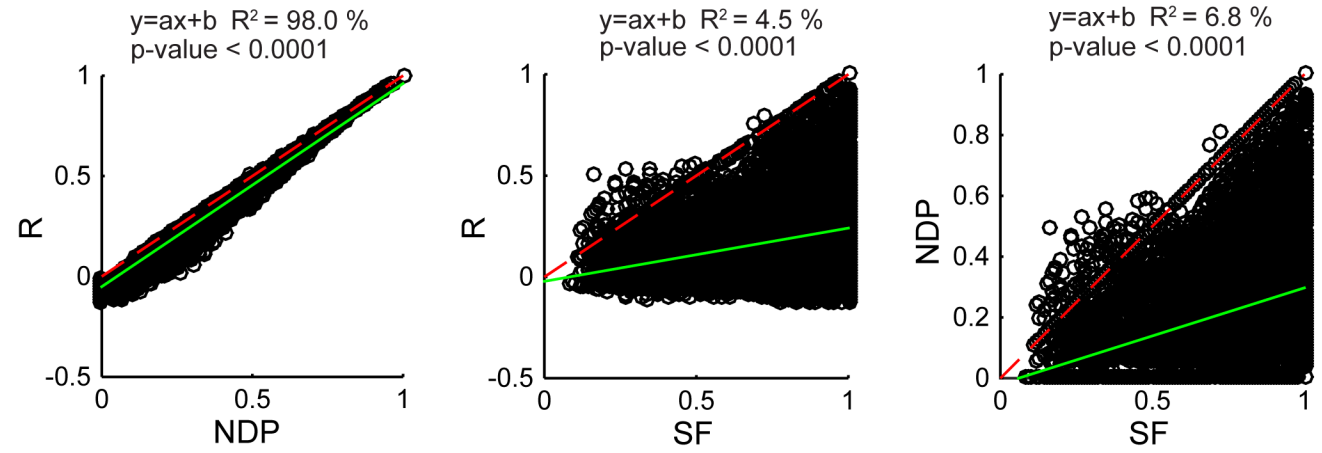
Figure 9



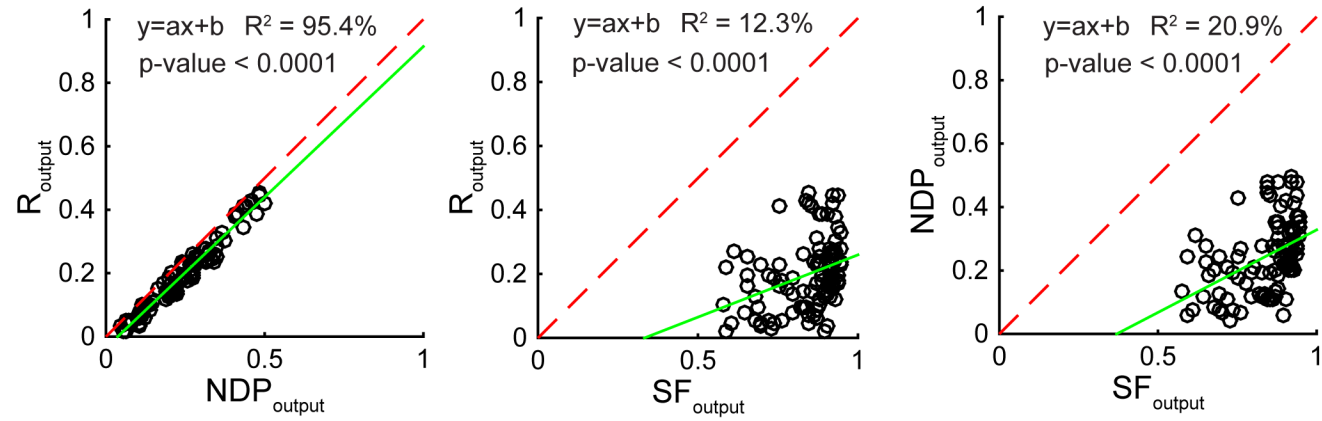
D Pairs of random synthetic spiketrains



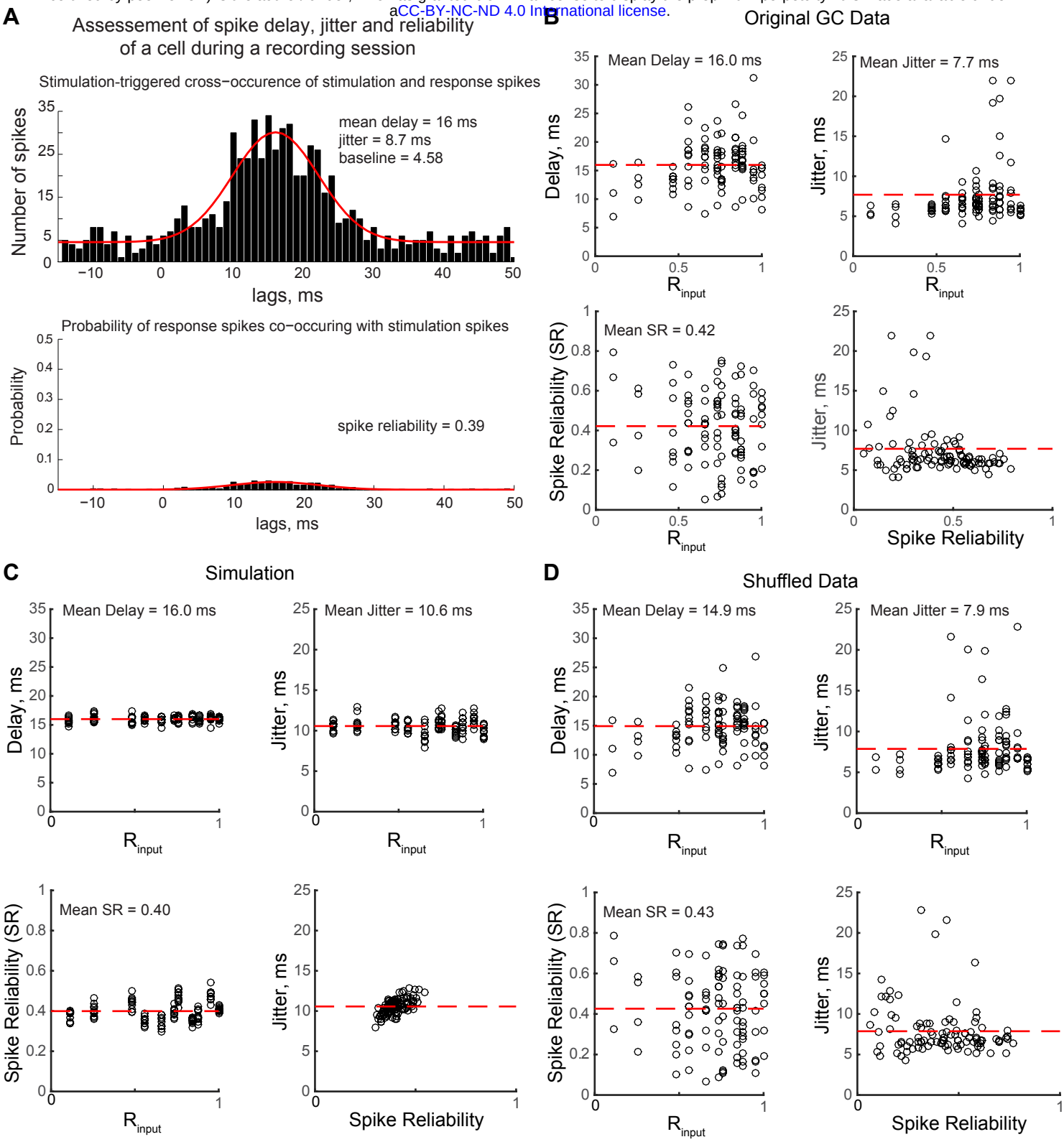
E Pairs of spiketrains from GC recordings



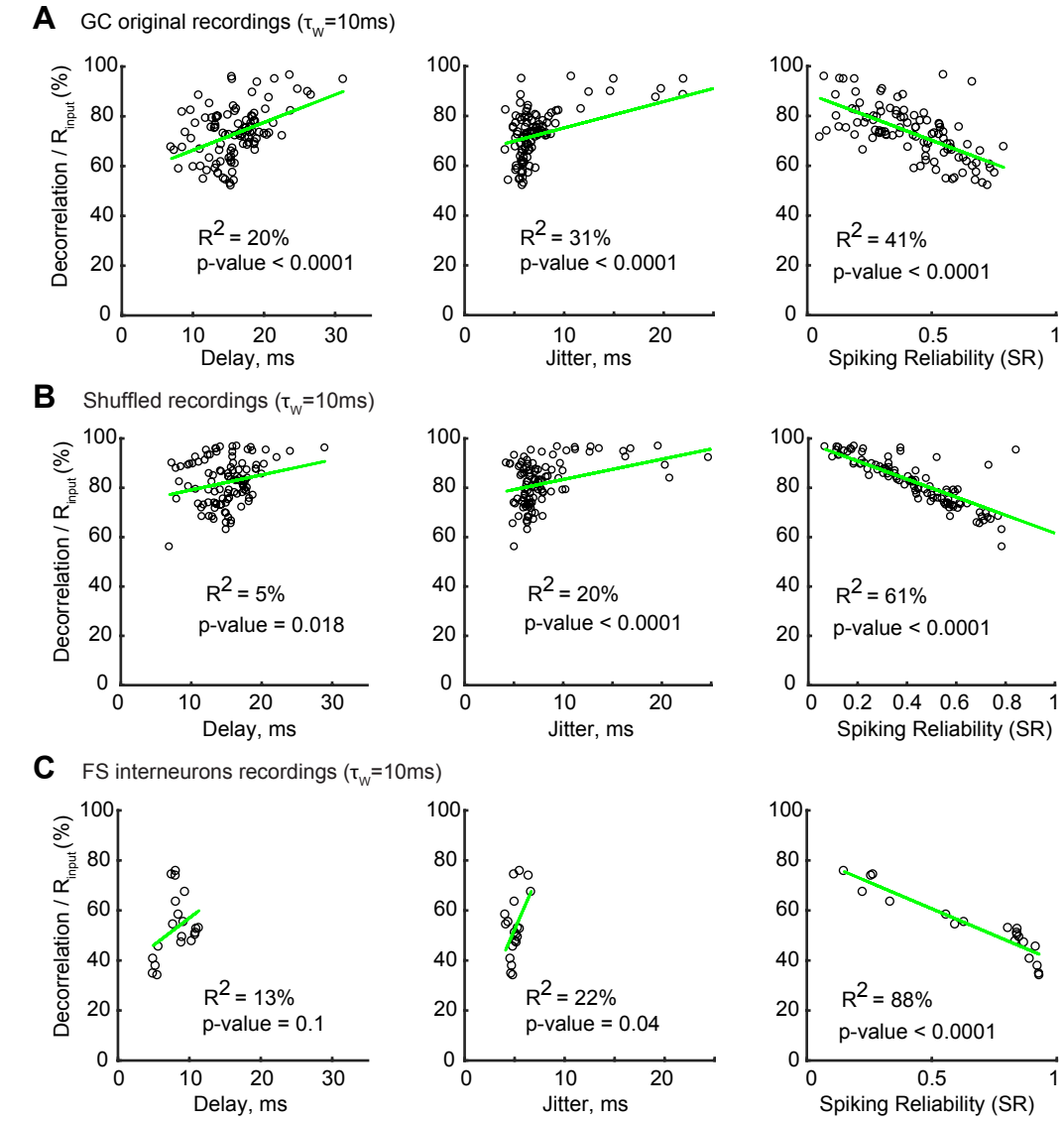
F Average per recording sets



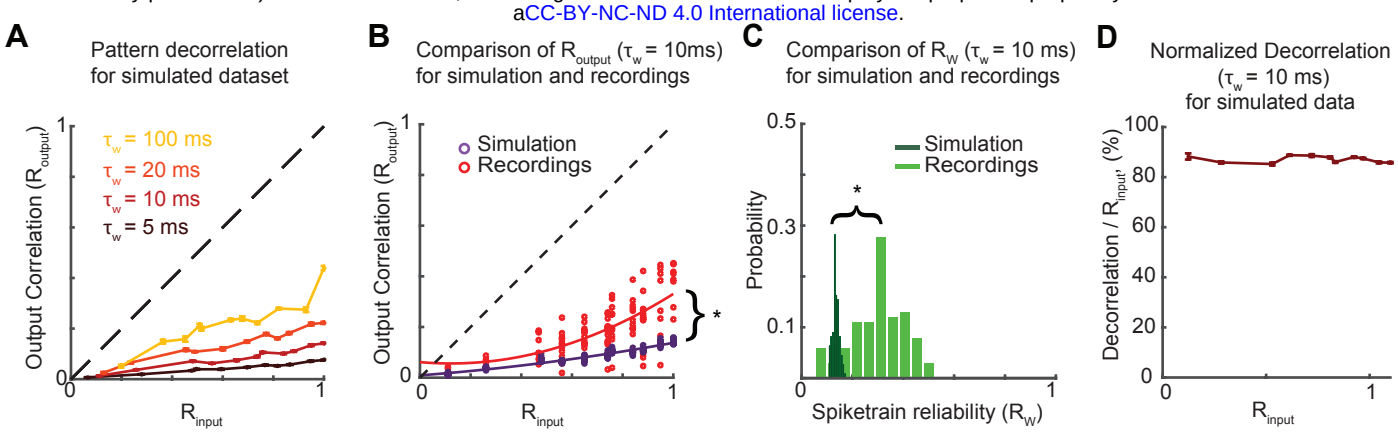
S1 Fig



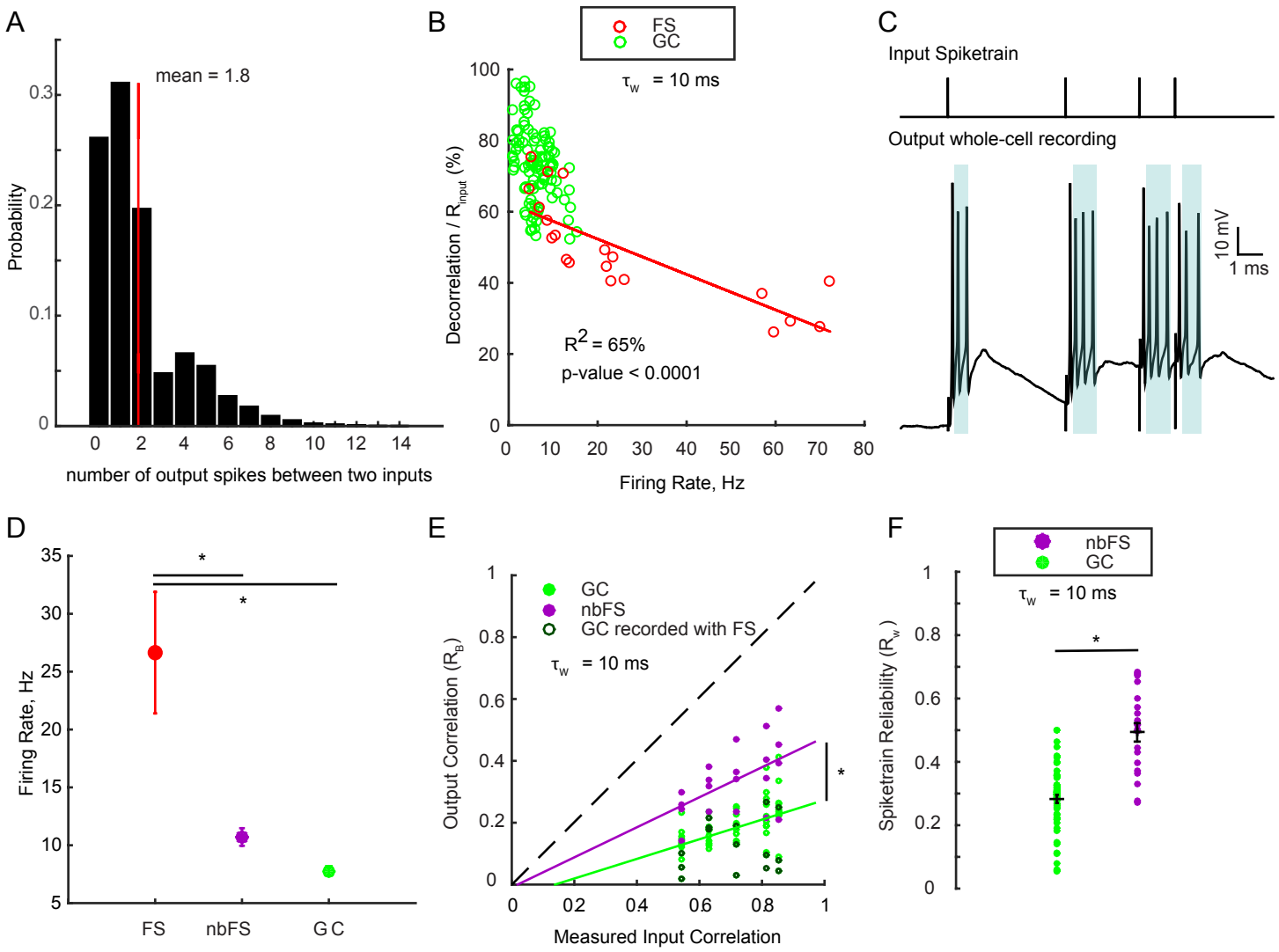
S2 Fig



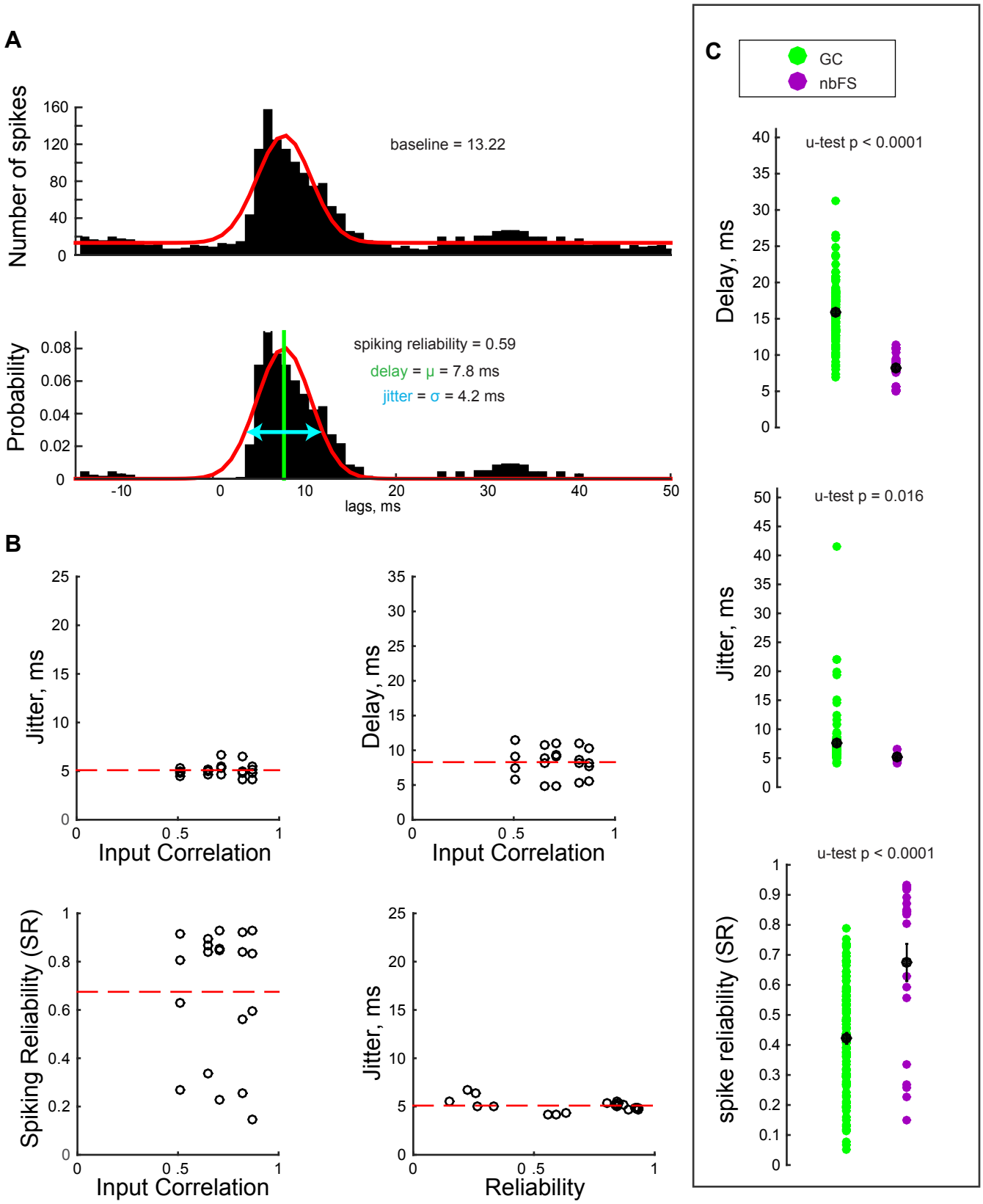
S3 Fig



S4 Fig



S5 Fig



S6 Fig

Table 1-3. **Linear regressions goodness-of-fit, p-value and slope.** The predictor variables (x-axis) correspond to columns, and the variables to be explained (y-axis) correspond to rows. Red highlights significant regressions that explain more than 50% of the variance ($R^2 > 50\%$). Blue highlights regressions that are significant ($p < 0.01$) but that explain less than 50% of the variance. The values used for Normalized Decorrelation, i.e. $(R_{input} - R_{output})/ R_{input}$, and for Spiketrain Reliability (Rw) were computed with a binning window of 10 ms, unless specified.

Table 1. Intrinsic electrophysiological cell properties

y-axis ↓ / x-axis →		**Membrane Capacitance (Cm)**	**Membrane Resistance (Rm)**	**Membrane Time Constant = Rm.Cm**	**Resting Membrane Potential (Vrest)**
Normalized Decorrelation	**GC**	$R^2 = 4\%$ $p = 0.08$ slope = -0.2	$R^2 = 5\%$ $p = 0.06$ slope = -0.03	$R^2 = 8\%$ $p = 0.013$ slope = -1.2	$R^2 = 3\%$ $p = 0.17$ slope = -0.2
	FS	$R^2 = 47\%$ $p = 0.0008$ slope = -1.5	$R^2 = 77\%$ $p < 0.0001$ slope = 0.6	$R^2 = 5\%$ $p = 0.4$ slope = 13.6	$R^2 = 46\%$ $p = 0.0009$ slope = 1.4
	GC + FS	$R^2 = 1\%$ $p = 0.3$ slope = -0.1	$R^2 = 4\%$ $p = 0.05$ slope = 0.04	$R^2 = 1\%$ $p = 0.4$ slope = -0.07	$R^2 = 0.1\%$ $p = 0.7$ slope = -0.1
Spiketrain Reliability (R_W)	**GC**	$R^2 = 2\%$ $p = 0.2$ slope = 0.001	$R^2 = 5\%$ $p = 0.03$ slope = 4e-3	$R^2 = 7\%$ $p = 0.02$ slope = 0.01	$R^2 = 3\%$ $p = 0.17$ slope = 0.002
	FS	$R^2 = 48\%$ $p = 0.0006$, slope = 0.01	$R^2 = 70\%$ $p < 0.0001$ slope = -0.007	$R^2 = 6\%$ $p = 0.29$ slope = -0.12	$R^2 = 39\%$ $p = 0.003$ slope = -0.014
	GC + FS	$R^2 = 0.7\%$ $p = 0.4$ slope = 0.001	$R^2 = 3\%$ $p = 0.07$ slope = -4e-5	$R^2 = 1\ \%$ $p = 0.37$ slope = 8e-4	$R^2 = 0.1\%$ $p = 0.7$ slope = 8e-4

Table 2. Spike-wise neural noise

y-axis ↓ \ x-axis →		**Delay**	**Jitter**	**Spiking Reliability (SR)**
Normalized Decorrelation	**GC**	$R^2 = 39.5\%$ $p < 0.0001$ slope = 1.8	$R^2 = 43\%$ $p < 0.0001$ slope = 1.3	$R^2 = 45\%$ $p < 0.0001$ slope = -47.3
	GC, 100ms	$R^2 = 2\%$ $p = 0.3$ slope = 0.7	$R^2 = 13\%$ $p = 0.01$ slope = 1.2	$R^2 = 41\%$ $p < 0.0001$ slope = -72
	FS	$R^2 = 13\%$ $p = 0.1$ slope = 2.2	$R^2 = 22\%$ $p = 0.04$ slope = 9.2	$R^2 = 88\%$ $p < 0.0001$ slope = -42.4
	Shuffle	$R^2 = 11\%$ $p = 0.02$ slope = 0.9	$R^2 = 24\%$ $p = 0.0003$ slope = 0.8	$R^2 = 65\%$ $p < 0.0001$ slope = -41.1
Spiketrain Reliability (R_W)	**GC**	$R^2 = 33\%$ $p < 0.0001$ slope = -0.016	$R^2 = 43\%$ $p < 0.0001$ slope = -0.013	$R^2 = 46\%$ $p < 0.0001$ slope = 0.5
	GC, 100ms	$R^2 = 3\%$ $p = 0.2$ slope = -0.008	$R^2 = 16\%$ $p = 0.004$ slope = -0.012	$R^2 = 43\%$ $p < 0.0001$ slope = 0.7
	FS	$R^2 = 15\%$ $p = 0.09$ slope = -0.024	$R^2 = 15\%$ $p = 0.09$ slope = -0.078	$R^2 = 85\%$ $p < 0.0001$ slope = 0.4
	Shuffle	$R^2 = 9\%$ $p = 0.04$ slope = -0.008	$R^2 = 23\%$ $p = 0.0004$ slope = -0.008	$R^2 = 65\%$ $p < 0.0001$ slope = 0.4

Table 3. Spiketrain-wise properties

y-axis ↓ \ x-axis →		**Overall Firing Rate**	**Spiketrain Reliability (R_W)**
Normalized Decorrelation	**GC**	$R^2 = 15\%$ $p < 0.0001$ slope = - 1.27	$R^2 = 81\%$ $p < 0.0001$ slope = -87
	FS	$R^2 = 65\%$ $p < 0.0001$ slope = - 0.50	$R^2 = 90\%$ $p < 0.0001$ slope = -85
	GC + FS	$R^2 = 47\%$ $p < 0.0001$ slope = - 0.80	$R^2 = 90\%$ $p < 0.0001$ slope = -90