the four groups form a 2 by 2 matrix, where groups on the same rows receive same amplitude of shock, and groups on the same column experience contexts in the same order. This design allow us to easily compare between groups, and uncover both the effects and interactions of emotional valence and temporal order.

In all groups, context A, B and C are distinct contexts that differ in lighting conditions, arena shapes, floor textures and scents. For "retrospective" experiments, animals are first put into context A to explore for 5 minutes. Then after a variable time intervals, the animals are put back into context B for 5 minutes, where they receive a delayed shock at fourth minute with an amplitude of either 0.75 mA or 1.5 mA, depending on whether the animals are in "low shock" or "high shock" group. The variable time intervals are achieved by further dividing animals into sub-groups according to different time intervals, and between-group comparison can be carried out. 2 days after these, the testing sequence is carried out, where the animals are put back to context A, B and C in that order, with 2d interval in between. Animals' freezing levels are assessed from behavior video recordings using standard software. For "prospective" experiments, the contexts, time intervals, shock intensity assignments and testing sequence remain identical to those in "retrospective" experiments. The difference is that in "prospective" experiments, the animals are first put into context B, where they receive a delayed shock, and then put into the neutral context A to explore.

According to preliminary results, we expect to see that in "low shock" and "retrospective" groups, animals are able to link context A and B together when they are separated by either 5 hours or 2 days, but not when they are separated by 4 days or 7 days. Specifically, within "low shock" and "retrospective" group, we expect to see higher freezing in context A for 5 hours and 2 days sub-group, but not for 4 days or 7 days sub-group. Similarly, we expect in "low shock" and "prospective" groups, the animals are only able to link together context A and B when they are separated by 5 hours, but not when they are separated by either 2 days, 4 days or 7 days. Such result would suggest that retrospective memory linking has longer temporal window than prospective memory linking. On the other hand, we expect to see in "high shock" and "retrospective" group, the animals may be able to link context A and B across either 5 hours, 2 days or 4 days, but not 7 days. Such result would suggest that more negative emotional value of a memory can extend the memory linking time window retrospectively.

Aim 2: Study how ensemble dynamic contribute to memory linking.

To study how ensemble dynamic contribute to memory linking, we will carry out calcium imaging studies shown in Figure 6. The design was similar to the behavior experiments in previous aim. Briefly, the animals are placed in context A to explore, then after a variable time interval, the animals are placed in context B, where they receive a delayed shock at second minute. Then two days later, the animals are put back in context A, context B, and a novel context C in that order for testing. Calcium imaging in behaving animals are carried out in all sessions. Thus the neural dynamics during all sessions can be collected and analyzed.

For simplicity, we use uppercase letters combined with numbers to denote the population of neurons that were active during a recording session: we use $A1$ and $A2$ to denote neurons that are active during the encoding and retrieval session of context A, respectively. We use $B1_{pre}$ and $B1_{post}$ to denote cells that active before and after the delivery of the shock during the encoding session of context B. In addition, we use $B2$ to denote active neurons in the retrieval session of context B, and lastly we use C to denote the active neurons encoding context C. Additionally, we use set operation conventions to denote the specific population of neurons that we are interested in. Specifically, we use set intersection "$\cap$"to denote the population of neurons that are active in both sessions. For example, $A1 \cap A2$ denote the cells that are active in both the encoding and the retrieval session of context A. Moreover, we use set difference "$\setminus$"to denote the population of neurons that are active in the first session but

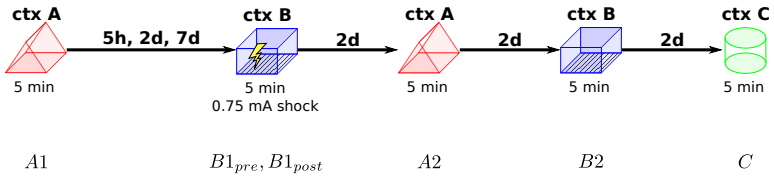

Figure 6: imaging experiment design

not the second session. For example, $A1 \setminus A2$ denote the cells that are active only during the encoding session of context A, but not reactivated during the retrieval session of context A. Finally, we use "$N\{\ldots\}$" to denote the number of cells in the population that we are interested in. For example, $N\{A1\}$ denote the total count of cells that are active during the encoding of context A.

To test whether representation of one context become more similar to the other after encoding, we first look at the reactivation rate of the neurons initially encoding a context during the retrieval of that context. Specifically, we calculate reactivation rate as the following:

$$R_{A1} = \frac{N\{A1 \cap A2\}}{N\{A1\}}$$
$$R_{B1pre} = \frac{N\{B1_{pre} \cap B2\}}{N\{B1_{pre}\}}$$
$$R_{B1post} = \frac{N\{B1_{post} \cap B2\}}{N\{B1_{post}\}}$$

where R_{A1}, R_{B1pre}, R_{B1post} denote the reactivation rate of the ensembles for context A, pre-shock phase for context B, and post-shock phase for context B, respectively. In addition, we also want to look at whether one context become more similar to the other after encoding. To test this, we can calculate a reactivation rate of the neurons initially encoding one context during the retrieval of the other context. Specifically, we calculate the rate as the following:

$$R_{BA} = \frac{N\{A1 \cap B2\}}{N\{A1\}}$$
$$R_{ABpre} = \frac{N\{B1_{pre} \cap A2\}}{N\{B1_{pre}\}}$$
$$R_{ABpost} = \frac{N\{B1_{post} \cap A2\}}{N\{B1_{post}\}}$$

Conceptually, R_{BA} indicate how close is the representation of context B during retrieval to the initial representation of context A. Similarly, R_{ABpre} and R_{ABpost}indicate how close is the representation of context A during retrieval to the initial pre-shock or post-shock representation of context B, respectively. If there is a drift of the representation of one context towards the other after encoding, we should see a decrease of reactivation rate of that context between encoding and retrieval, as well as a higher overlap between the retrieval of that context and the encoding of the other context.

To study how different ensembles during encoding may contribute to the overlap during the retrieval, we can calculate the proportion of cells during encoding that also serve as the overlapping cells during retrieval. Specifically:

$$C_X = \frac{N\{X \cap (A2 \cap B2) \setminus (A2 \cap B2 \cap C)\}}{N\{X\}}$$

where X is one of the three ensembles during encoding, that is: $A1$, $B1_{pre}$, or $B1_{post}$. The term $(A2 \cap B2) \setminus (A2 \cap B2 \cap C)$ give us the population of neurons that are active specifically in both the retrieval of context A and B, but not active during all three retrieval sessions of context A, B and C. The intersection between this population and the ensembles of one of the encoding sessions will then estimate the contribution of neurons from one of the encoding sessions to this specific population that may drive memory linking. Thus, C_{A1}, C_{B1pre} and C_{B1post} will indicate the contributions to the overlap during retrieval from the ensembles encoding context A, pre-shock phase of context B, and post-shock phase of context B respectively. We can then compare these three rate to see whether different ensembles during encoding contribute equally to the population of neurons that may drive memory linking.

Lastly, in line with the excitability hypothesis, we expect the activity level of a neuron should affect the likelihood of that neuron being reactivated during future sessions. Specifically, for each neuron, the activity level of that

neuron can be summarized by the area-under-the-curve of the calcium trace of that neuron. By taking a z-score of such activity level across all neurons in a recording session, we can obtained a centered activity level for each neuron that's also normalized to the overall activity level of all neurons in that session. We can denote such centered and normalized activity level in session X as $A_X\{\ldots\}$, where $\ldots$ indicate a specific population we are interested in. Then, for any given two recording session X and Y, where X happens before Y, four quantities can be calculated and compared:

$$\begin{aligned} Activity_X^{common} &= A_X\{X \cap Y\} \\ Activity_X^{alone} &= A_X\{X \setminus (X \cap Y)\} \\ Activity_Y^{common} &= A_Y\{X \cap Y\} \\ Activity_Y^{alone} &= A_Y\{Y \setminus (X \cap Y)\} \end{aligned}$$

$Activity_X^{common}$ indicate the activity level in X of those cells that were reactivated in Y, whereas $Activity_X^{alone}$ indicate the activity levels in X of the cells that were only activated during X, but not Y. We can then compare the mean of $Activity_X^{common}$ and $Activity_X^{alone}$. According to excitability hypothesis, we expect to see higher $Activity_X^{common}$ than $Activity_X^{alone}$, since the activity level in X should affect whether they are reactivated in later sessions. On the other hand, we expect no difference between $Activity_Y^{common}$ and $Activity_Y^{alone}$, since the activity level of cells in the later session should not affect whether they are co-activated in both sessions.

Aim 3: Develop analysis pipeline for calcium imaging with miniature microscope.

To study the neural correlate of memory linking, we would carry out miniature calcium imaging in behaving animals. We choose miniature calcium imaging due to its capability to record neuronal activities in behaving mice and to track same field-of-view across long period of time, which is essential for the purpose of memory linking studies. We will focus on recording in dorsal CA1 region since it is believed to make a major contribution to contextual memory. The experimental design is identical to Aim 1 as shown in Figure 5. Neuronal activities would be recorded during all experiment sessions to allow collection of rich dataset.

The raw videos from calcium imaging recording could be processed with an open-source analysis toolkit CaImAn implementing a constrained non-negative matrix factorization algorithm. After the process, a spatial matrix representing the spatial footprint of each putative neurons, as well as a temporal matrix representing the calcium traces of each putative neurons will be extracted from the raw data. A custom-written script is used to visually assess the accuracy of the extraction as well as manually refine the results. After this, neurons from different recording sessions are cross-registered based on the euclidean distances between the centroids of their spatial footprint, and a unique master index can be assigned to each neuron in the whole experiment.

The main comparison is the ensemble overlap between context A and B, B and C, as well as A and C during retrieval/testing. We expect to see a higher overlap between context A and B when the two context are linked together and animals exhibit elevated freezing level in both of them, while the overlap between A and B are not expected to be significantly higher than those between B and C or A and C when the two contexts are not linked together.

In addition, a PCA analysis can be carried out to reveal temporal structures of each ensemble. Specifically, for each recording session, given a matrix representing the calcium traces of N neurons along T time-steps (usually frames), a PCA analysis can be applied to extract R principal components, where each components contain a "neuron vector" $\vec{n}$ of length N, and a "temporal vector" $\vec{t}$ of length T. Thus the dimension of the data is reduced from $N \times T$ to $R \times (N + T)$. The PCA is carried out in a way so that: a) the "neuron vector" of each principle component represent a group of neurons that has a highly correlated firing pattern, and the "temporal vector" represent that averaged pattern treating the whole group as single neuron. b) a dot product can be computed with each "neuron vector" and "temporal vector", and the sum of R such dot products should closely reproduce the original $N \times T$ data. c) the R components should explain most of the variance in the original data, thus the value of R can be determined by thresholding the proportion of variance explained.

Once the principal components of each recording sessions are extracted, we can calculate a cross-correlation of the "neuron vector"s between any two session. We can then compare such correlation matrices between linked