Detection of the spike times

Figure 1 displays, as examples, typical intensity time traces at low and high pump current (26.5 mA and 27.3 mA for the experimental data and $\mu = 0.98$ and 1.01 for the numerical data). For both, the following method was used to detect the times when the spikes occur. First, each time series was normalized to have zero mean value and standard deviation equal to one. Next, the spike times were determined by using a simple thresholding method: when the intensity decreases below a certain value, a spike is recorded. Then, it was required that the intensity grows above zero before another spike can occur. This avoids detecting spurious spikes, specially during the intensity recovery (the dropout is sharp but the recovery is gradual and noisy). Alternatively, one could use a “refractory time” after each dropout, during which the intensity is recovering and no spikes are detected.

To verify the method of spike detection we compare the mean inter-spike interval, $\langle \Delta T \rangle$, and the normalized standard deviation, $\sigma / \langle \Delta T \rangle$, for experimental and simulated data. Figure 2 displays $\langle \Delta T \rangle$ and $\sigma / \langle \Delta T \rangle$ vs. the pump current (experimental data), and vs. the parameter $\mu$ (numerical data), calculated using a
spike detection threshold equal to -2. A good qualitative agreement experiments-simulations can be observed, and the variation of \( \langle \Delta T \rangle \) and \( \sigma / \langle \Delta T \rangle \) with the pump current is also in good agreement with previous observations [1, 2].

To further verify the method of spike time detection, we calculated the histograms of \( \Delta T \) values. Figure 3 displays the histograms corresponding to the time series shown in Fig. 1, again computed with a threshold of -2. We observe a good qualitative agreement of the experimental and numerical histograms, which are also in good agreement with previous reports [3, 4]. We note that for low pump current, in the numerical histogram (Fig. 3b) there is a sharp peak at low \( \Delta T \) values, that is not seen in the experimental histogram (Fig. 3a). The reason of this peak is the presence of intermittent bursts of regular spikes. This regular dynamics has been observed experimentally in the literature and has been referred to as regular pulse packages [5, 6]. Because the regular spikes are not as deep as the irregular ones (Fig. 4a), by choosing a deeper threshold (e.g., -2.5) they are not detected and the peak in the histogram of \( \Delta T \) values is eliminated, as seen in Fig. 4b.

**Influence of the detection threshold in the word probabilities**

Next, we analyze the influence of the detection threshold in the probabilities of the ordinal patterns. Figure 5 shows the probabilities vs. the detection threshold, for the same pump currents as in Fig. 1. It can be appreciated that, while the values of the probabilities vary with the threshold, the hierarchy and the clusters (021-102 and 120-201) are robust and occur in a wide range of threshold values. While for too low (or too deep) thresholds the probabilities vary significantly (as too many or too few spikes are detected), they are robust to threshold variations in a wide range of thresholds.

Most importantly, the variation of the probabilities with the threshold is qualitatively the same in the experimental and in the numerical data. The same hierarchy and clusters are seen. This is remarkable because the model used for the simulations is the simplest rate-equation model (assumes single-mode emission, neglects spatial and thermal effects, considers only optical noise, etc.) and the filter used to simulate the finite detection bandwidth is also a simple moving-average window.

While an optimal threshold could be defined for each pump current value (that is in the center of the “plateau” where the probabilities do not vary significantly with the threshold), for the sake of simplicity in our work we used a fixed threshold value, equal to -2, for detecting the spikes.
Figure 1: Typical time series of the laser intensity. Panels a and c display experimental data (the pump current is 26.5 mA and 27.3 mA respectively) and panels b and d, numerical data (the pump current parameter is $\mu = 0.98$ and 1.01 respectively). The dots indicate the times when a spike is detected using a threshold equal to -2. In the simulations $\eta = 20$ ns$^{-1}$ and $\alpha = 4.5$, other parameters are as indicated in Methods.
Figure 2: (a, b) Mean inter-spike interval and (c, d) normalized standard deviation versus the pump current, for experimental data (a, c) and for numerical data (b, d). The threshold for detecting the spike times is -2 and the parameters of the simulation are as in Fig. 1.

Figure 3: Histograms of the inter-spike intervals corresponding to the time series shown in Fig. 1.
Figure 4: (a) Detail of the regular pulse packages in the simulated time-series for $\mu = 0.98$. The dots indicate two thresholds for detecting the spike times. Panel (b) displays the histogram of $\Delta T$ values computed by using the lower threshold (-2.5).

**Calculation of the error bars of probability values and the null hypothesis region.**

In order to determine if the word probabilities are consistent with a distribution of equally probable words (null hypothesis, N.H., of no correlations present in the spike sequence), we calculate the error bars of the probabilities as well as the probability region consistent with the N.H. They are calculated as in Ref. [7]: using a binomial test to compute the error bars, and using a confidence level of 95% for the N.H. region: $p \pm 3\sigma_p$, where $p = 1/D!$ and $\sigma_p = \sqrt{(p(1-p)/N)}$, with $D$ being the length of the word and $N$ the number of words in the sequence.

Figures 6a-c display the results for words of length $D = 2, 3$ and 4 respectively. One can clearly observe that the distribution of probability values is not consistent with the N.H., as several probabilities are outside the N.H. region (gray region in Fig. 6). For the sake of clarity, Fig. 6d displays the probabilities of words of length $D = 4$ computed from surrogated (shuffled data), and we can confirm that in this case the probabilities are all within the gray region.

**Analysis of spike correlations**

To further confirm the presence of correlations in the spike sequence, we compute the usual autocorrelation coefficient of the sequence of inter-spike intervals, $\{\Delta T_i\}$, as
Figure 5: Probabilities of the words vs. the detection threshold is shown for different pump current values both, for the experimental (a, c) and the simulations (b, d) data.
Figure 6: Probabilities of the words of length (a) $D = 2$, (b) $D = 3$ and (c) $D = 4$ vs. the pump current, computed from the experimental data. The words '012' and '210' present the same crossover as words '01' and '10', and '0123' and '3210' (blue and red, respectively in panel c) as words '012' and '210'. Panel (d) displays the probabilities for $D = 4$ computed from the surrogated (shuffled) data.
Figure 7: First and second order correlation coefficient of the inter-spike-interval sequence, $C_1$ and $C_2$ [Eq. 1], computed from the experimental data, as a function of the laser pump current.

$$C_\tau = \frac{\langle (\Delta T_i - \langle \Delta T \rangle)(\Delta T_{i-\tau} - \langle \Delta T \rangle) \rangle}{\sigma^2}$$  \hspace{1cm} (1)

The result is shown in Fig. 7, that displays $C_1$ and $C_2$ vs. the pump current for the experimental data. First-order correlations are clearly identified, which are stronger at low pump currents, while second-order correlations are significantly weaker.

From these results we can also confirm that the symbolic ordinal analysis indeed provides additional information with respect to the usual correlation analysis. In particular, we notice that $C_1$ is positive for all current values; however, in Fig. 6a we see that at high currents the word '10' is more probable (i.e., $\Delta T_i > \Delta T_{i+1}$ is more probable than $\Delta T_i < \Delta T_{i+1}$), and at low currents, it is the opposite situation (the word '01' is more probable, and thus, $\Delta T_i < \Delta T_{i+1}$ is more probable than $\Delta T_i > \Delta T_{i+1}$). This is apparently contradictory; however, one should keep in mind that the ordinal analysis takes into account the relative values of consecutive intervals, while the correlation coefficient, takes into account the magnitude of consecutive $\Delta T_i - \Delta T >$ values. Moreover, Fig. 7 shows that $C_2$ is very small, suggesting negligible correlations; however the probabilities of the words of length $D = 3$ are clearly not consistent with the N.H.
Additional experimental measurements

In order to demonstrate the robustness of our observations, we performed several experiments employing different lasers and feedback conditions. Figures 8(a) and 8(b) display the probabilities computed with data analyzed in our previous work [7]. In that experiment we used a different laser (Hitachi Laser Diode HL 6724 MG), lasing at 675 nm, and the feedback was stronger (the threshold reduction being 7%). Two data sets were recorded at different temperatures (18°C in panel a, 20°C in panel b). One can observe that in both data sets, the hierarchy, clusters and crossover are the same as in Fig. 6b. For Fig. 8c the strength of the optical feedback is weaker, resulting in a threshold reduction of 4%. Also the two clusters and a crossover (such that the word 210 becomes the most probable one at high current) are seen in this data set.

Also additional measurements were done including external forcing via direct modulation of the laser pump current. A 1550 nm semiconductor laser (Mitsubishi ML925B45F) was used, with feedback provided by an optical fiber. The feedback conditions were significantly different from those of the experiment reported in the main text. In this new experiment the feedback time delay was longer (25 ns), and the feedback stronger (threshold reduction being 11%). The dc value of the pump current was 12.50 mA (1.12I_{th}) and temperature was 17°C. Figure 8d displays the probabilities vs. dc pump current modulation amplitude and we can observe that, in spite of the fact that the experimental conditions are very different, there is a good qualitative agreement with Fig. 4 in the main text. For strong enough modulation amplitude, the hierarchy and clusters are the same in both data sets.

References


Figure 8: Words probabilities computed from data recorded using different experimental conditions. Panels (a) and (b) present the analysis of the data of ref. [7], recorded for two laser temperatures, (a) 18 C and (b) 20 C. The threshold reduction was 7% and the time delay 3 ns. The laser used was different from that referred to in the main text. The data analyzed in panel (c) was recorded using the same laser as in the main text, but with a different feedback level (threshold reduction of 4%). Time delay was 4.7 ns. The data analyzed in panel (d) correspond to a periodically modulated laser, emitting at telecom wavelength (1550 nm) and with optical feedback provided via an optical fiber that results in a threshold reduction of 11%. The feedback time delay was 25 ns and the laser was pumped at a dc current of 12.50 mA (equal to 1.12$I_{th}$).

