Supplementary Information: Temporal pattern recognition through analog molecular computation

Jackson O'Brien and Arvind Murugan

Methods

We first formulated abstract chemical networks with the desired feature detecting properties. Then, largely following the design principles laid out in,[1] we implemented these chemical reaction networks as DNA strand displacement reactions. All strand displacement circuits are designed within Microsoft's Visual DSD software described in[2] using default kinetic parameters and concentrations $\in [.05 \text{ nM}, 10\ \mu\text{M}]$. Then, by adapting the MATLAB code generated within this program, we exposed these circuits to the pulsatile inputs defined in the main text, defined by their duty fraction δ, number of pulses n, and period T. All results shown are from deterministic simulations without leak reactions.

Detailed Chemical Networks

All reactions shown utilize the default kinetic parameters within the Visual DSD software (3×10^{-4} $\frac{1}{\text{nM s}}$ bind, .1226 $\frac{1}{\text{s}}$ unbind corresponding to toe-holds with 4-6 nucleotides[3]). Different reaction rates were achieved by selecting appropriate initial concentrations. The dynamic range we require ($\sim$ 5 orders of magnitude) is also achievable through toehold design.[3] Species with specified initial concentrations are outlined in bold and their values are given in accompanying tables. All species whose initial concentrations are specified and do not have explicit time dependence displayed in the main text are held at their initial concentrations throughout all simulations. In the tables of concentrations below, parenthetical values indicate high and low oscillatory values in the time-varying input.

Pulse Counting

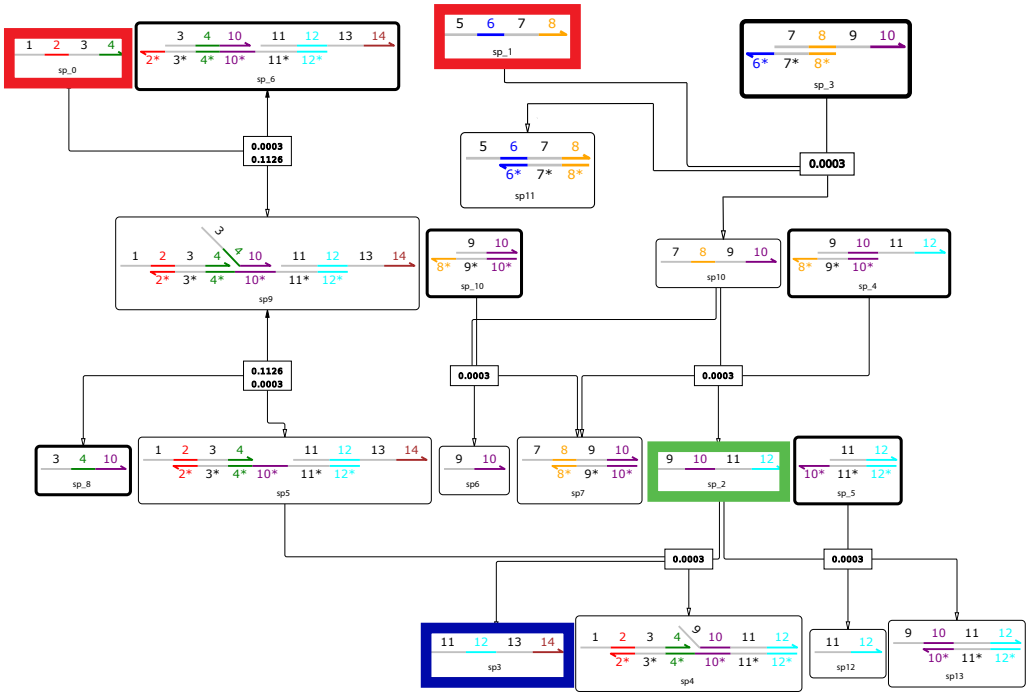

Figure 1: Full version of DNA strand displacement reaction network for pulse counting decoder including waste products. The two species in red had their dynamics directly modulated to the parameters of the input series, with sp_1 (referred to as $\bar{I}$ in the main text) pulsing exactly out of phase with sp_0 (I in the main text). Graphs and labels are generated automatically within the Visual DSD software.[2] See Table 1 for a list of initial concentrations. Note that the flux G plotted in Figure 2 is defined $G \equiv (\sqrt{\mathrm{sp}_2 * \mathrm{sp}_5})$.

Table 1: Pulse Counter Initial Conditions

Species	Initial Conc. (nM)
sp_0	1 (0)
sp_1	100 (0)
sp_2	50
sp_3	10000
sp_4	10000
sp_5	10
sp_6	10000
sp_8	10000
sp_10	10000

Duty Fraction

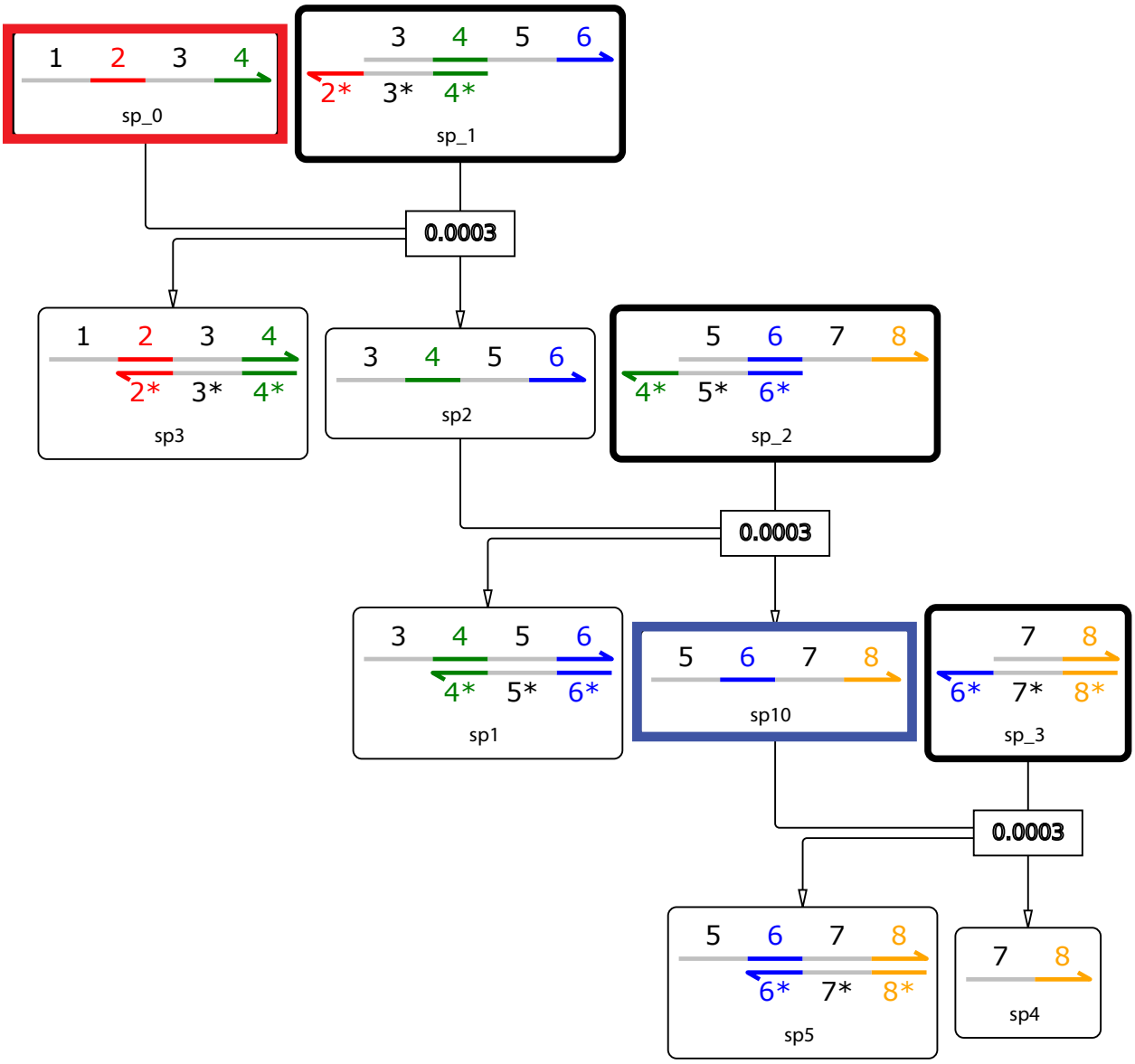

Figure 2: Full version of DNA strand displacement reaction network for duty cycle decoding, including waste products. This circuit effectively takes the moving average of the dynamics of sp_0 and reports it in sp10. See main text Figure 3 for analysis and Figure 5 for performance. Initial concentrations of bolded species are listed in Table 2.

Table 2: Duty Fraction Decoder Initial Conditions

Species	Initial Conc. (nM)
sp_0	1(0)
sp_1	100
sp_2	100
sp_3	10

Period Detecting

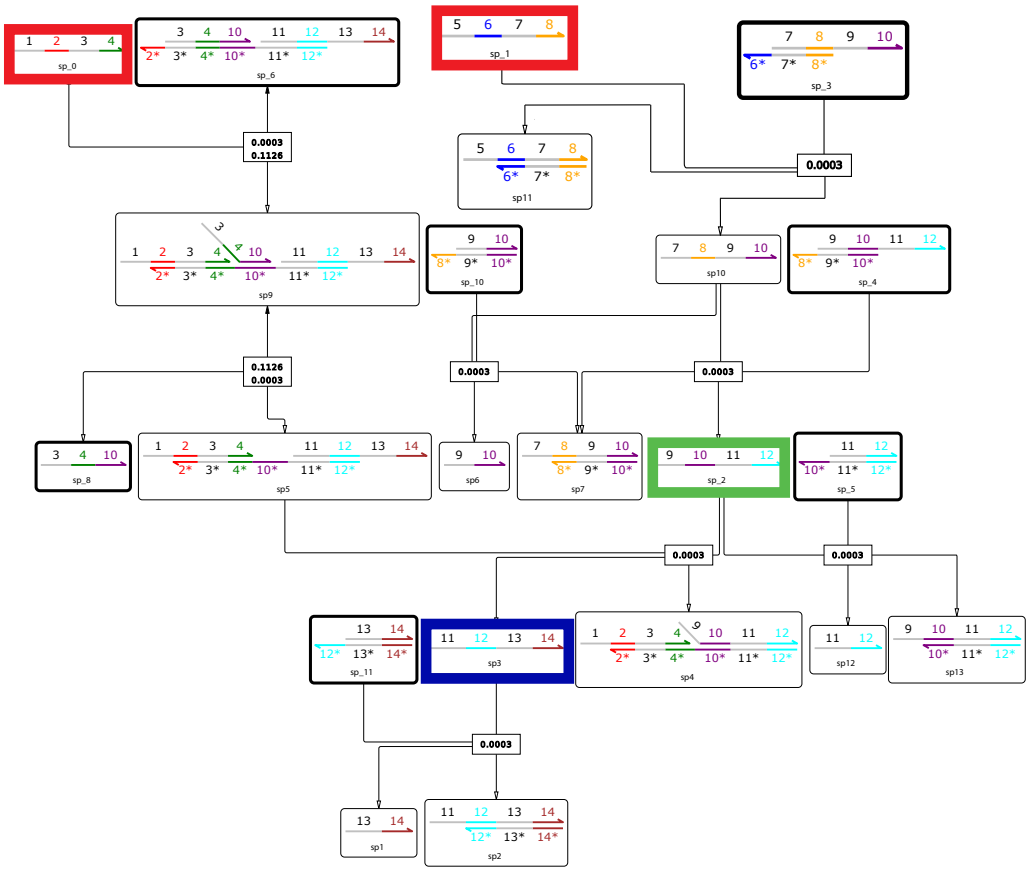

Figure 3: Full version of DNA strand displacement reaction network for time period decoding, including waste products. By taking the moving average of the incoming flux $G \equiv \sqrt{(\mathrm{sp}_2 * \mathrm{sp}_5)}$, this circuit decodes the period of sp_0. Initial concentrations of bolded species are reported in Table 3.

Table 3: Period Decoder Initial Conditions

Species	Initial Conc. (nM)
sp_0	1 (0)
sp_1	100 (0)
sp_2	50
sp_3	10000
sp_4	10000
sp_5	10
sp_6	10000
sp_8	10000
sp_10	10000
sp_11	.1375

Supplemental Calculations

Pulse Counter

$$\mathrm{I} + \mathrm{A} \xrightarrow{k_1} \mathrm{P} \tag{1}$$

$$\mathrm{A} \xrightarrow{\lambda_1} \phi \tag{2}$$

$$\bar{\mathrm{I}} \xrightarrow{k_2} \mathrm{A} \tag{3}$$

$$\tag{4}$$

To supplement the analysis presented in the main text, we present the system of corresponding ordinary differential equations governing the evolution of A and P, presented in Equation 6.

$$\dot{A}(t) = k_2\bar{I}(t) - k_1A(t)I(t) - \lambda_1A(t) \tag{5}$$

$$\dot{P}(t) = k_1A(t)I(t) \tag{6}$$

The discussion of timescales relies on exponential decays and exponential approaches to steady state exhibited by these differential equation in transitioning between the two states $I = C, \bar{I} = 0$ and $I = 0, \bar{I} = C$ where C corresponds to the finite value presented in Table 1 for sp_0 and sp10.

When I is turned on, P starts being produced since P requires both I and A to be present (Note that A has a resting state at a high concentration $\frac{k_2 C}{\lambda_1}$ when the input I is off). However, turning on I also causes exponential decay of A from $\frac{k_2 C}{\lambda_1}$ to 0 on a timescale $\frac{1}{k_1 C + \lambda_1}$.Consequently, the term promoting P decays after a short transient. Thanks to the simple form of these equations, we can compute $P(t)$ analytically

$$P(t) = \frac{C^2 k_1 k_2}{\lambda_1} \frac{(1 - e^{-(\lambda_1 + k_1 C)t})}{k_1 C + \lambda_1} \tag{7}$$

which exhibits the dependence of P on the degradation timescale of A, $\tau_{\mathrm{a}} = \frac{1}{\lambda_1 + k_1 C}$.

Note that any pulse of width,

$$T\delta \gg \tau_{\mathrm{a}}$$

will produce a stereotyped exponential decay profile (i.e. independent of T and δ) for P. This restriction provides one limit on the window of operation for pulse counters.

Finally, once I switches off, A returns to its steady state $\frac{k_2 C}{\lambda_1}$ at a timescale $\frac{1}{\lambda_1}$. The pulse needs to be off for long enough

$$T(1 - \delta) \gg \frac{1}{\lambda_1},$$

so that A can be restored to its original state between every pulse. This condition ensures each step up of P will have the same value (the ODE for P only depends on the values of I and A) and can serve as a proxy for pulse number n.

Within the regime defined by the two inequalities, we find that the circuit presented is able to count pulses independent of duty fraction or time period.

Duty Fraction Decoder

The duty fraction decoder is given by the simple network,

$$\mathrm{I} \xrightarrow{k} \mathrm{I} + \mathrm{P} \tag{8}$$

$$\mathrm{P} \xrightarrow{\lambda} \phi \tag{9}$$

and consequently, $P(t)$ is governed by,

$$\dot{P}(t) = kc(t) - \lambda P(t) \tag{10}$$

where the input concentration $I = c(t)$ is taken to vary as $c(t)$. The solution to the above equation can be written in terms of a exponentially decaying kernel,

$$P(t) \;=\; k \int_{-\infty}^{t} c(t') e^{-\lambda(t-t')} dt'. \tag{11}$$

P(t) will be at its smallest value at the front edge of a pulse; this value, after $n-1$ pulses can be computed from the above,

$$P(t = (n-1)T) \;=\; \frac{k}{\lambda} \frac{(e^{T\lambda\delta} - 1)(1 - e^{-(n-1)T\lambda})}{e^{T\lambda} - 1} \tag{12}$$

P(t) will be largest right at the end of the pulse; this value, after n pulses can be computed to be,

$$P(t = (n-1+\delta)T) \;=\; \frac{k}{\lambda} \frac{(e^{T\lambda\delta} - 1)(e^{(1-\delta)\lambda T} - e^{-(n-1+\delta)T\lambda})}{e^{T\lambda} - 1}. \tag{13}$$

We consider the average of these two quantities as representing the readout value of P.

Adding and simplifying yields

$$\bar{P} \;= \left(\frac{e^{\lambda T\delta}-1}{2(e^{\lambda T}-1)}\right)\; (1 + e^{\lambda T(1-\delta)} - e^{-\lambda T(n-1)} - e^{-\lambda T(\delta+(n-1))}). \tag{14}$$

In the limit of

$$n \gg 1/(\lambda T), \quad T \ll 1/\lambda,$$

$\bar{P} \approx \frac{k\delta}{\lambda}$ is proportional to the duty fraction δ but independent of n and T.

Finally, we can consider the difference between the maximal and minimum values of $P(t)$ as a measure of the variation away from this readout. Taking the difference of Equations 13 and 12 yields

$$\Delta P \;= \left(\frac{e^{\lambda T\delta}-1}{(e^{\lambda T}-1)}\right)\; (-1 + e^{\lambda T(1-\delta)} + e^{-\lambda T(n-1)} - e^{-\lambda T(\delta+(n-1))}). \tag{15}$$

For large n and small λT, this approximates to $\Delta P \approx \frac{k\delta}{\lambda}\lambda T(1-\delta) \approx \bar{P}\lambda T(1-\delta)$. Thus, in the limit defined above, $T \ll \frac{1}{\lambda}$, the fractional variation about $\bar{P}$ defined as $\frac{\Delta P}{\bar{P}}$ is small, and $\bar{P}$ is a reliable readout for the duty fraction δ.

References

(1) David Soloveichik, E. W., Georg Seelig *Proceedings of the National Academy of Sciences* **2010**, *107*, 5393–5398.

(2) Lakin, M. R.; Youssef, S.; Polo, F.; Emmott, S.; Phillips, A. *Bioinformatics* **2011**, *27*, 3211–3213.

(3) Zhang, D. Y.; Winfree, E. *Journal of the American Chemical Society* **2009**, *131*, 17303–17314, PMID: 19894722.