Nonequilibrium associative retrieval of multiple stored self-assembly targets: Supporting information

Gili Bisker[a,b,1] **and Jeremy L. England**[a,1]

[a]Physics of Living Systems Group, Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [b]Department of Biomedical Engineering, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel

Multiple-target storage challenge

The number of target structures that can be encoded and retrieved with high fidelity using the specific interactions between distinguishable building blocks is limited. We follow the logic of (1) and consider N particles with m possible target states, where each target consists of a specific spatial arrangement of the N particles. Particles that are nearest neighbors according to one or more of the targets, strongly attract. In this setting, each particle can potentially bind to Zm partners, where Z is the valency, or the number of nearest neighbors. Hence, the probability that two particles specifically attract is Zm/N. For a particle to stably bind to a growing seed, we assume it needs to form $Z/2$ strong bonds on the seed's boundary. The number of particles that can thus bind to an available boundary site is $N(Zm/N)^{Z/2}$, whose floor value (rounding down to the nearest integer) should equal one for a high fidelity retrieval. Hence, the number of targets that can be stored and retrieved, solely based on geometrical considerations, is $\lfloor N^{(Z-2)/Z}/Z \rfloor$.

Let us stress that this bound does not take into account thermodynamic considerations, which are introduced through the parameters of the system such as the temperature T and the strong specific interaction energy J_s. For a ratio J_s/T too high, the system would be stuck in kinetic traps and local minima, whereas a value too low would result in a homogeneous phase above the melting point such that no nucleation would occur. Hence, the bound can be realized given that the system's parameters are carefully tuned.

In our case, we introduce another layer of complexity by allowing particles to switch their internal state, and in turn, their inter-particle interaction energies. According to our model, in order for two particles to experience strong attraction, they need to be nearest neighbors in one of the encoded targets, and adopt the appropriate internal states, according to that target. Otherwise, the particle pair would experience weak or intermediate attraction, as detailed in the main text. However, the scaling behavior of the limit of the number of targets remains similar to the above derivation, and poses a challenge on storing too many targets.

The R Factor

The factor R, which appears in the main text in Eq. 8 can be further simplified using the parameters of the system. Let us denote $D_\alpha \equiv J_\mathrm{w} - J_\mathrm{s}^\alpha$ for targets $\alpha = 1$ and $\alpha = 2$, respectively. Then,

$$\begin{cases} E_\alpha^c - E_\alpha = D_\alpha \pm h_1 \\ E_\alpha^e - E_\alpha = 1.5D_\alpha \pm h_1 \\ E_\alpha^b - E_\alpha = 2D_\alpha \pm h_1 \end{cases} \quad [S1]$$

corresponding to a single corner particle flip, an edge particle flip, and a bulk particle flip, respectively, and the $-$ and $+$ signs correspond to targets $\alpha = 1$ and $\alpha = 2$, respectively. Note that we have used the fact that in our model $h_2 = 0$. Further,

$$\begin{cases} E_\alpha^{*c} - E_\alpha = -2J_\mathrm{s}^\alpha \\ E_\alpha^{*e} - E_\alpha = -3J_\mathrm{s}^\alpha \end{cases} \quad [S2]$$

corresponding to a corner particle and an edge particle detaching from the target cluster, respectively. Combining all of the above and plugging in d_c, d_e, and d_b, we find:

$$R \equiv \frac{e^{h_1}[4e^{-D_1} + 4(\sqrt{N}-2)e^{-1.5D_1} + (\sqrt{N}-2)^2 e^{-2D_1}] + 8e^{2J_\mathrm{s}^1} + 4(\sqrt{N}-2)e^{3J_\mathrm{s}^1}}{e^{-h_1}[4e^{-D_2} + 4(\sqrt{N}-2)e^{-1.5D_2} + (\sqrt{N}-2)^2 e^{-2D_2}] + 8e^{2J_\mathrm{s}^2} + 4(\sqrt{N}-2)e^{3J_\mathrm{s}^2}} \quad [S3]$$

Supporting Figures

This section includes figures with additional numerical data of the multi-target self-assembly process. In Fig. S1A the first assembly time is plotted as a function of the driving force for different number of targets, whereas in Fig. S1B we plot the total entropy production until a target is assembled normalized by the value of the drive, as a function of the driving force for

[1]G.B. and J.L.E. contributed equally to this work.

different number of targets. In the simulations included in this figure, the initial positions of the particles on the lattice were randomly chosen. In contrast, for the results presented in Fig. S2, we initialized the system at one of the target states. We plot the time spent at the target state, and the mean distance from target along the trajectory as a function of the driving force.

In Fig. S3, we plot the energy and the distance from the targets for the case of 2 encoded targets, with a bias toward target 1 applied by a stronger binding interaction, as a function of time, with and without the drive. The bias towards target 1 is obvious from the data, as the system spends more time closer to target 1 along the simulation trajectory. With the added drive, the energy values along the simulation are lower, and the system spends more time closer to both targets, compared to the un-driven realization.

Fig. S4 shows the energy levels spacing between the microstates corresponding to the assembled targets, and the microstates corresponding to a single particle flipping its internal state or detaching from the target cluster.

In Fig. S5, we illustrate the deviation from the Boltzmann ratio by plotting the ratio between the time spent at target 2 and the time spent at target 1, for the two cases of bias considered in the work, different internal energies and different binding energies. In Fig. S6, we plot the saturation value of the deviation as a function of the R factor. The fit coefficients for both cases of the bias shown in Fig. S5, as well as the combined case discussed in the main text, collapse onto the same curve.

Finally, in Fig. S7, we show the scaling of the R factor with the total number of particles, and in Fig. S8 we plot the energy difference between the first excited state and the ground state, as a function of the internal energy bias and the binding energy bias for both targets, to illustrate the difference between the two cases.

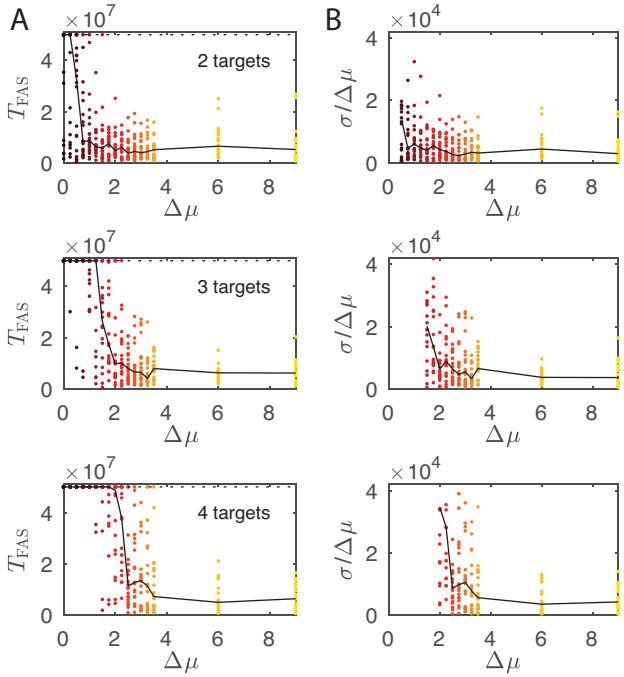

Fig. S1. (A) First assembly time, T_{FAS}, of one of the encoded targets as a function of the driving force value for 2, 3, and 4 targets (top to bottom, respectively). The black curve is the median assembly time as a unction of the driving force value, to provide a guide to the eye. (B) The total entropy production up to the first assembly time, $\sigma/\Delta\mu$, normalized by the the driving force value as a function of the driving force value. The black curve is the median value of $\sigma/\Delta\mu$ as a unction of the driving force value, to provide a guide to the eye.

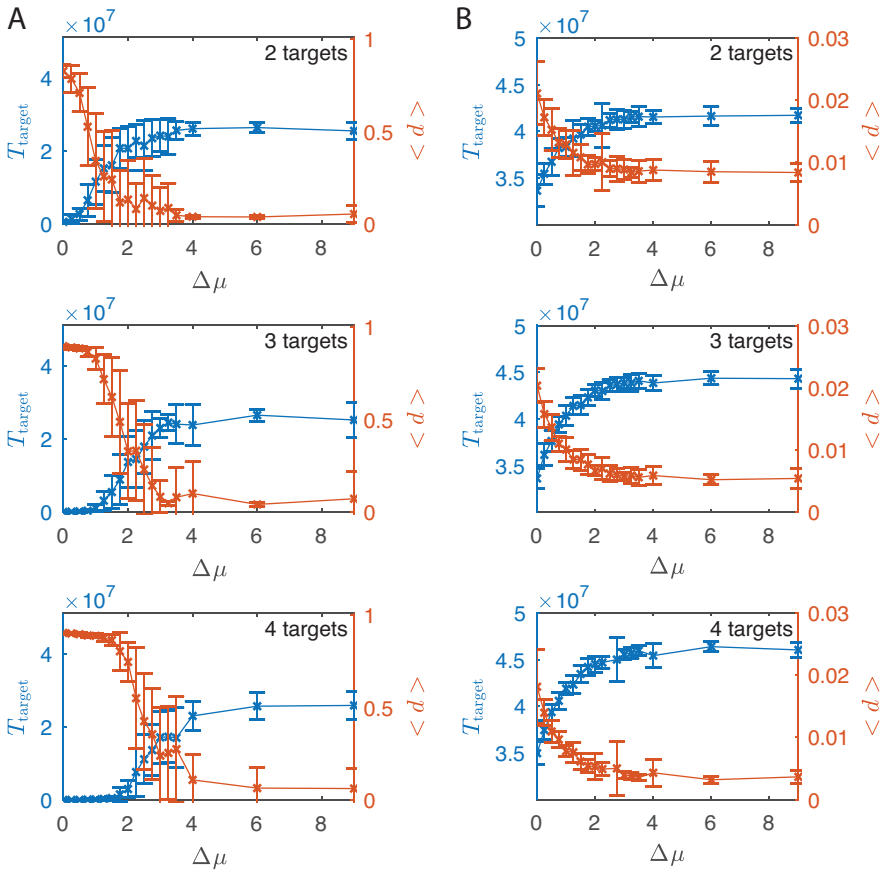

Fig. S2. Time spent at the target, T_{target}, for simulations initialized at one of the target structures (blue crosses, left y-axis), and the mean distance from the target, $< d >$, (orange crosses, right y-axis), as a function of the driving force value $\Delta\mu$ for two (top), three (middle), and four (bottom) possible targets. (A) Binding interaction $J_{\text{s}} = -4$ corresponding to the fastest assembly when a single target is encoded. (B) Binding interactions for the fastest assembly when multiple targets are encoded: $J_{\text{s}} = -4.6$, $J_{\text{s}} = -4.8$, and $J_{\text{s}} = -5$, for two, three, and four targets, respectively. The data points are averages over 20 realizations with identical initial conditions of target 1. Error bars are of one standard deviation.

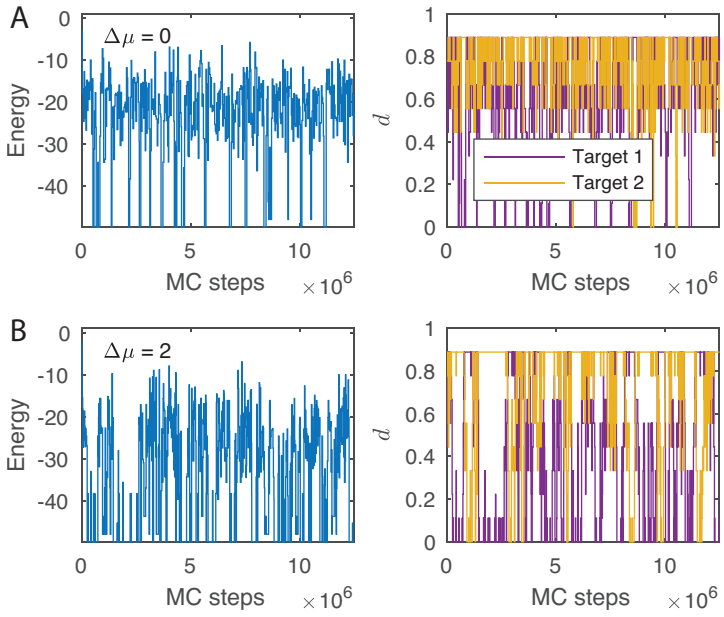

Fig. S3. Target bias. (A) Typical realization of the simulation for $m = 2$ targets, where target 1 is favored over target 2 by stronger binding energy, $J_{\text{s}}^1 = -4.15$, and $J_{\text{s}}^2 = -4$, respectively, either in detailed balance, $\Delta\mu = 0$, or (B) with the local drive $\Delta\mu = 2$. The plots in the left column show the energy along the simulation as a function of the Monte Carlo (MC) step number. The plots on the right column show the distance d from target 1 (purple curve) and target 2 (yellow curve) as a function of the MC step number. For an assembled target, $d = 0$.

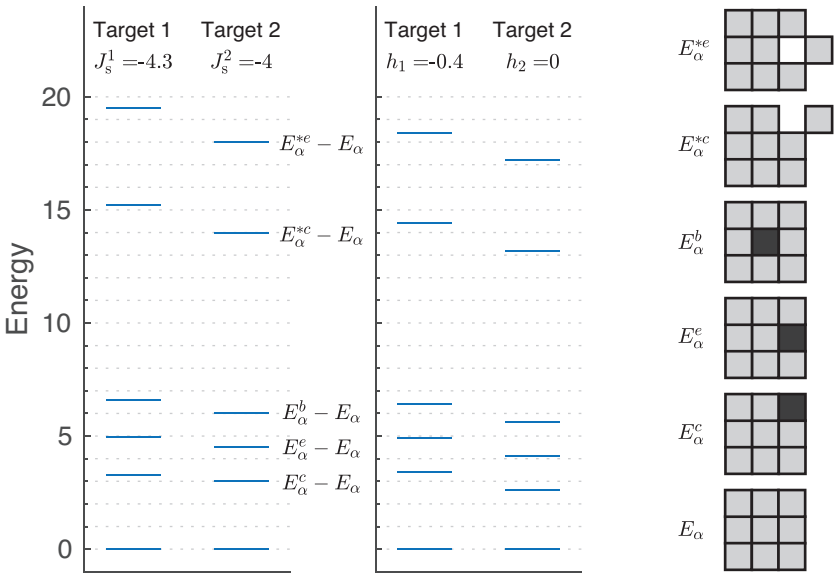

Fig. S4. Energy levels spacing. Energy levels for targets 1 and 2 relative to the respective ground states for different binding energies (left panel) and different internal energies (middle panel). The lowest energy levels represent the state where all the particles are at the target structures, with the same corresponding internal configuration (represented by the light gray fill color in the right panel). The excited states correspond to a single corner-particle flip (represented by the dark gray fill color), a single edge-particle flip, a single bulk-particle flip, a single corner-particle detaching, and a single edge-particle detaching, respectively (right panel).

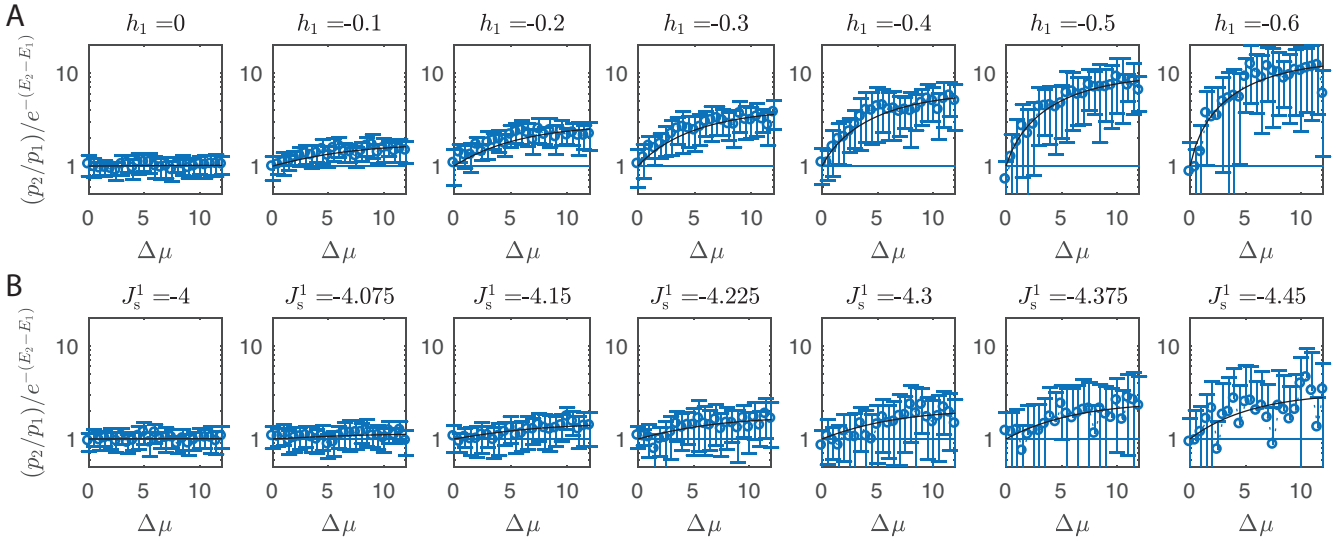

Fig. S5. Deviation from the Boltzmann ratio. (A) The ratio between the time spent at target 2 and the time spent at target 1, p_2/p_1,normalized by the corresponding Boltzmann ratio, $e^{-(E_2-E_1)}$, as a function of the drive value $\Delta\mu$ for different internal energies h_1 while $h_2 = 0$ is held constant, and (B) different binding energies J_s^1 while $J_s^2 = -4$ is held constant. The solid horizontal blue curve is the baseline, 1. The solid black line is a fit to the data according to $c_1 \tanh(c_2\Delta\mu) + 1$.

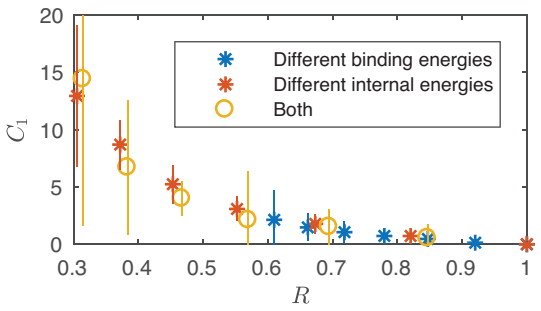

Fig. S6. The deviation from the Boltzmann ratio at saturation, c_1, as a function of R. The vertical lines correspond to the 95% confidence intervals of the fit.

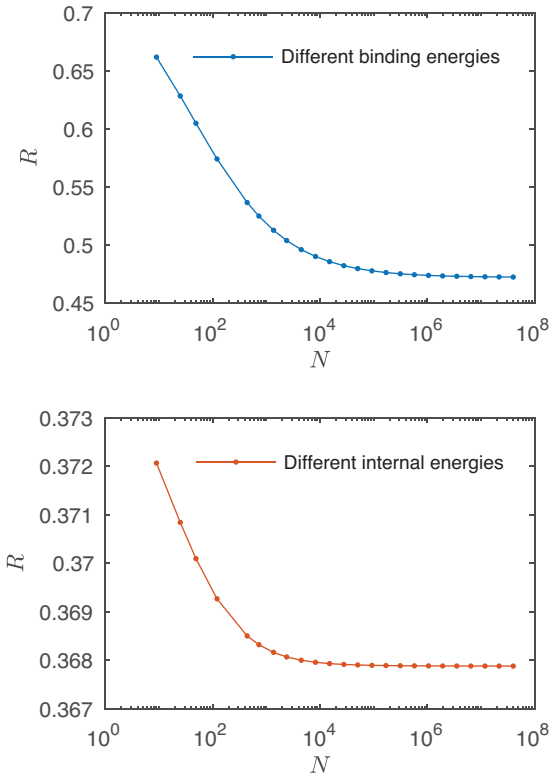

Fig. S7. Scaling of R as a function of N for binding energy bias (top panel), where $J_{\mathrm{s}}^{1} = -4.45$, and for internal energy bias (bottom panel), where $h_1 = -0.6$.

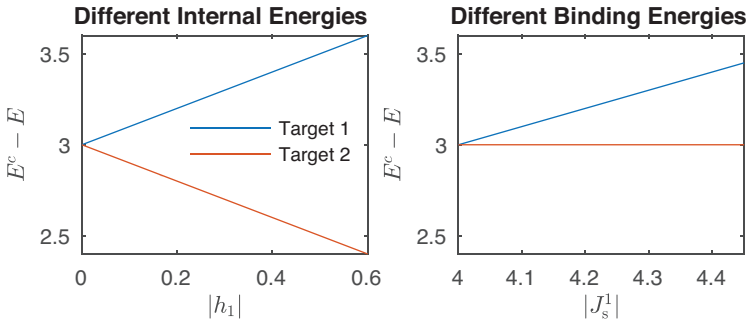

Fig. S8. The energy difference between the first excited state and the ground state, $E_{\alpha}^{c} - E_{\alpha}$ as a function of $|h_1|$ (left panel) and $|J_{\mathrm{s}}^{1}|$ (right panel), for target 1 and 2.

1. Murugan A, Zeravcic Z, Brenner MP, Leibler S (2015) Multifarious assembly mixtures: Systems allowing retrieval of diverse stored structures. *Proceedings of the National Academy of Sciences* 112(1):54–59.