

1

2 **Supplementary Information for**
3 **Non-equilibrium correlations in minimal dynamical models of polymer copying**

4 **Jenny M Poulton, Pieter Rein ten Wolde and Thomas E Ouldridge**

5 **Thomas E Ouldridge.**

6 **E-mail: t.ouldridge@imperial.ac.uk**

7 **This PDF file includes:**

8 Supplementary text

9 Figs. S1 to S2

10 References for SI reference citations

11 Supporting Information Text

Solving for partial velocities and tip probabilities. Following ref. (1), the partial velocities can also be expressed as

$$v_x = \frac{\psi_{xr}^+ v_r}{\psi_{xr}^- + v_r} + \frac{\psi_{xw}^+ v_w}{\psi_{xw}^- + v_w}, \quad [1]$$

for $x = r, w$. This gives a pair of simultaneous equations that can be solved for the partial velocities in terms of the propensities. In turn, the velocities and propensities determine tip and conditional probabilities $\mu(m_i)$ and $\mu(m_{i-1}|m_i)$, with $m_i, m_{i-1} \in \{r, w\}$, via

$$\mu(x) = \frac{\psi_{rx}^+}{\psi_{rx}^- + v_x} \mu(r) + \frac{\psi_{wx}^+}{\psi_{wx}^- + v_x} \mu(w), \quad [2]$$

$$\mu(x|y) = \frac{\psi_{xy}^+}{\psi_{xy}^- + v_y} \frac{\mu(x)}{\mu(y)}. \quad [3]$$

12 **Difference between the tip probabilities and the final sequence.** It might not be immediately obvious why the properties of the
 13 growing chain described by $\mu(m_i)$ and $\mu(m_i, m_{i-1})$ should be different from those of the final chain described by ϵ , ϵ_r and ϵ_w ,
 14 but the difference can be illustrated with a simple example. Consider a system in which incorrect monomers could be added to
 15 the end of the chain, but where nothing can be added after an incorrect match. In this case while the tip probability for an
 16 incorrect match $\mu(w)$ would be finite, the error of the final chain ϵ would be vanishingly small, as all incorrect matches would
 17 have to be removed in order for the polymer to grow further.

18 **Demonstrating that the sequence of the final chain is Markovian.** Let M_n^∞ be the monomer at the n th site in the final chain.
 19 Let a polymer be represented by $M_1^\infty, \dots, M_n^\infty$. The probability of a given chain existing is then $P^\infty(m_1, \dots, m_n)$. In order for
 20 the sequence of monomers moving along the chain (increasing n) to be able to be represented by a Markov chain, the condition

$$P^\infty(m_n|m_{n-1}, \dots, m_1) = P^\infty(m_n|m_{n-1}) \quad [4]$$

22 must hold.

23 In order to demonstrate that eq. 4 holds, we rewrite the final chain probability in terms of properties of the growing chain.
 24 Specifically we state that the probability of the sequence m_1, \dots, m_n existing in the final chain is the product of the probability
 25 $Q(m_1, \dots, m_{n-1}, t)$ that the chain is in the state m_1, \dots, m_{n-1} at a time t during the growth process, and the propensity
 26 $\nu(m_n; m_1, \dots, m_{n-1})$ with which m_n is added to a chain m_1, \dots, m_{n-1} and never removed, integrated over all time. It should be
 27 noted that $r(m_n; m_1, \dots, m_{n-1})$ is time-independent.

$$P^\infty(m_1, \dots, m_n) = \int Q(m_1, \dots, m_{n-1}, t) \nu(m_n; m_1, \dots, m_{n-1}) dt \quad [5]$$

$$P^\infty(m_1, \dots, m_n) = \nu(m_n; m_1, \dots, m_{n-1}) \int Q(m_1, \dots, m_{n-1}, t) dt \quad [6]$$

31 Setting the integral equal to $I(m_1, \dots, m_{n-1})$ gives

$$P^\infty(m_1, \dots, m_n) = \nu(m_n; m_1, \dots, m_{n-1}) I(m_1, \dots, m_{n-1}) \quad [7]$$

33 Let's consider the probabilities of two sub-sequences, identical except for the final monomer. We call the two final monomers
 34 m_n and m'_n and we can denote the ratios of the probabilities of these two chains as follows.

$$\frac{P^\infty(m_1, \dots, m_n)}{P^\infty(m_1, \dots, m'_n)} = \frac{\nu(m_n; m_1, \dots, m_{n-1}) I(m_1, \dots, m_{n-1})}{\nu(m'_n; m_1, \dots, m_{n-1}) I(m_1, \dots, m_{n-1})} \quad [8]$$

36 The I terms are independent of this final monomer and so cancel. Thus

$$\frac{P^\infty(m_1, \dots, m_n)}{P^\infty(m_1, \dots, m'_n)} = \frac{\nu(m_n; m_1, \dots, m_{n-1})}{\nu(m'_n; m_1, \dots, m_{n-1})} \quad [9]$$

38 The same relationship holds for the conditional probabilities

$$\frac{P^\infty(m_n|m_1, \dots, m_{n-1})}{P^\infty(m'_n|m_1, \dots, m_{n-1})} = \frac{r(m_n; m_1, \dots, m_{n-1})}{r(m'_n; m_1, \dots, m_{n-1})} \quad [10]$$

40 For our system, the propensity with which a monomer m_n is added and never removed, $\nu(m_n; m_1, \dots, m_{n-1})$, is dependent on
 41 only on the final two monomers in the sequence fragment, m_n and m_{n-1} . To see why, note that this propensity is determined
 42 by addition and removal of monomers at sites $i \geq n$. The identities of monomers at positions $j < n - 1$, however, only influence
 43 addition and removal propensities at sites $k < n$ (eq. 1-4 main text). Thus we convert $\nu(m_n; m_1, \dots, m_{n-1})$ to $f(m_n, m_{n-1})$.

$$\frac{P^\infty(m_n|m_1, \dots, m_{n-1})}{P^\infty(m'_n|m_1, \dots, m_{n-1})} = \frac{f(m_n, m_{n-1})}{f(m'_n, m_{n-1})} \quad [11]$$

45 Multiplying both sides by $P^\infty(m_1, \dots, m_{n-2})$ and summing over all values of m_1, \dots, m_{n-2} (recall $\sum_{c,d} P(a|b, c, d)P(c, d) =$
 46 $P(a|b)$ and $\sum_{c,d} P(c, d) = 1$) gives

$$47 \frac{P(m_n|m_{n-1})}{P(m'_n|m_{n-1})} = \frac{f(m_n, m_{n-1})}{f(m'_n, m_{n-1})}. \quad [12]$$

48 Comparing equations 11 and 12 yields:

$$49 \frac{P^\infty(m_n|m_{n-1}, \dots, m_1)}{P(m'_n|m_{n-1}, \dots, m_1)} = \frac{P^\infty(m_n|m_{n-1})}{P^\infty(m'_n|m_{n-1})}. \quad [13]$$

50 Summing over the possible values of m'_n and recalling that $P^\infty(r|m_{n-1}) + P^\infty(w|m_{n-1}) = 1$ and $P^\infty(r|m_{n-1}, \dots, m_1) +$
 51 $P^\infty(w|m_{n-1}, \dots, m_1) = 1$ yields:

$$52 P^\infty(m_n|m_{n-1}, \dots, m_1) = P^\infty(m_n|m_{n-1}), \quad [14]$$

53 thereby proving that the sequence of the final chain is Markovian.

Overall error probability of the final chain. The final sequence is described by the Markov chain illustrated in figure S1. The transition matrix for this process is

$$T = \begin{bmatrix} 1 - \epsilon_r & 1 - \epsilon_w \\ \epsilon_r & \epsilon_w \end{bmatrix}, \quad [15]$$

The eigenvector of this transition matrix with eigenvalue equal to unity gives the steady state of the Markov chain. The second component of this eigenvector corresponds to the overall probability of incorrect matches, ϵ :

$$\epsilon = \epsilon_r / (1 + \epsilon_r - \epsilon_w). \quad [16]$$

54 **Corroboration with simulation.** To check the analytical methods used to solve the system we also simulated the growth of a
 55 polymer. We used a Gillespie simulation (2), with transition rates given by ψ_{xy}^\pm . Simulations were initialised with a randomly
 56 determined two monomer sequence, and truncated as soon as the polymer reached 1000 monomers. We found that such a
 57 length rendered edge effects negligible in all but the most extreme cases for the calculation of ϵ . Polymer error probabilities
 58 were inferred directly from the 100 simulations, and are compared to analytical results in fig. S2.

59 We note in passing that the calculation of H , H_{ss} , and particularly the efficiency's η and η_{ss} , are more vulnerable to random
 60 fluctuations in a simulation of finite length, and peculiar edge effects, than ϵ . Gaspard's solution is therefore invaluable in
 61 reaching robust conclusions for these quantities.

62 References

- 63 1. Gaspard P, Andrieux D (2014) Kinetics and thermodynamics of first-order markov chain copolymerization. *J CHEM PHYS*
 64 141:044908.
- 65 2. Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. *J CHEM PHYS* 81:2340–2361.

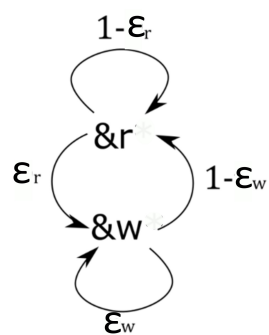


Fig. S1. The transition diagram for the Markov process describing the sequence of monomers found by stepping forward along a completed chain.

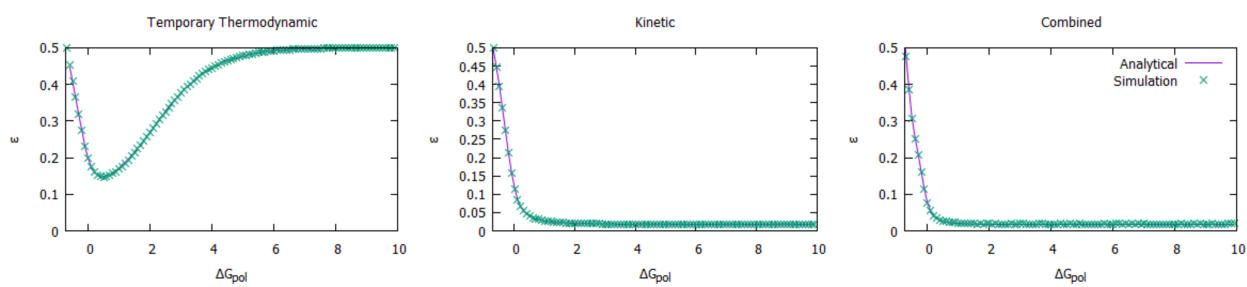


Fig. S2. Errors obtained from Gillespie simulations are indistinguishable from the analytical results obtained using Gaspard's method for all three mechanisms.