Appendix

December 20, 2019

1 Motor–bundle system approximation

Although the *N*-dimer long motor-bundle system described in the main text is a finite-dimensional Markov chain (in which each state denotes a unique combination of each MT's length and state, as well as the motor's conformational state), it contains too many possible states, and has too complicated of a topology, to make the analysis of the arrival times tractable. Instead, we present several physical arguments for reducing the system's complexity by introducing position-dependent effective rates, numerically verify the similarity between the full and reduced systems, and then analyze the reduced system analytically.

Physically, the most important parameters in simulating a motor's trajectory along a bundle is its initial position along the bundle axis, and whether it is in a diffusing state, attached to a MT growing towards the right, or attached to a MT growing towards the left. We therefore attempt to simplify our full model with the simplified motor-bundle model in Figure 1, in which each state denotes a motor's position and state with respect to the bundle.

Instead of considering all the conformational changes of the motor, we introduce an effective stepping rate parallel to the MT, $\lambda_{\uparrow\uparrow}$, an effective stepping rate antiparallel to the MT, $\lambda_{\uparrow\downarrow}$, and a detachment rate, μ . To take into account the walkoff effect described in the main body of this work, we denote by f_i the probability that a motor attached *i* dimers away from a MT's nucleation site walks off the tip upon stepping parallel to the MT; thus, a motor attached at the *i*th dimer of a right-growing MT detached at a rate of $\mu + f_i \lambda_{\uparrow\uparrow}$, and moves to the *i* + 1th dimer at the reduced rate $(1 - f_i)\lambda_{\uparrow\uparrow}$. From the diffusing state, a motor reattaches to some MT at a rate R_i , with the probability p_i that it's a right-growing MT. We derive these rates below.

2 Effective stepping and reattachment rates

For the coarser Markov chain model described above, we need to find an effective rate at which a motor moves from one fully attached state, to either of the two adjacent attached states, or to a diffusing state.

In Figure 2 the 11 state denotes a motor fully attached to a MT; the 01 state denotes the same motor with the head closer to the minus end of the MT detached; and the 10 state denotes the motor with the head closer to the plus end of the MT detached. The + state denotes the motor moving one full dimer parallel to the MT axis, the D state denotes the motor detaching from the MT completely, and the – state denotes the motor moving one full dimer antiparallel to the MT axis. Thus the excepted time τ for a motor to switch between one fully attached state to an adjacent fully attached state, or a fully detached state, may be obtained by treating the +, D, – states as absorbing, and computing the expected absorption time from

Figure 1: A simplified Markov Chain describing the movement of a motor along a MT bundle. A motor can be in one of three states: attached to a MT growing towards the right, diffusing, or attached to a MT growing towards the left.



Figure 2: Conformational change in motor stepping one along one dimer forward or back.

the 11 state. This is given by

$$\tau = -(1,0,0) \begin{bmatrix} -(\beta_b + \beta_f) & \beta_b & \omega \\ (1-p)\omega & -(\omega + \overline{\beta}) & 0 \\ p\omega & 0 & -(\omega + \overline{\beta}) \end{bmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$
(1)

$$=\frac{\overline{\beta}+\beta_b+\beta_f+\omega}{\overline{\beta}\beta_b+p\overline{\beta}\omega+\overline{\beta}\beta_f+(1-p)\beta_f\omega+\beta_f\omega}.$$
(2)

We therefore assign the effective rates as

$$\lambda_{\uparrow\uparrow} = \frac{p_+}{\tau},\tag{3}$$

$$\mu = \frac{p_D}{\tau},\tag{4}$$

$$\lambda_{1\downarrow} = \frac{p_{-}}{\tau},\tag{5}$$

where p_+ , p_D and p_- are the corresponding probabilities of absorption starting from the 11 state, which we obtain from the embedded chain of Figure 1.

In particular, let

$$A = \begin{bmatrix} 0 & \frac{\beta_b}{\beta_b + \beta_f} & \frac{\beta_f}{\beta_b + \beta_f} \\ \frac{\omega(1-p)}{\overline{\beta} + \omega} & 0 & 0 \\ \frac{\omega p}{\overline{\beta} + \omega} & 0 & 0 \end{bmatrix}$$
(6)

be the matrix denoting the transition probabilities between the transient states of the embedded chain, and

$$R = \begin{bmatrix} 0 & 0 & 0\\ \frac{\omega p}{\overline{\beta} + \omega} & \frac{\overline{\beta}}{\overline{\beta} + \omega} & 0\\ 0 & \frac{\overline{\beta}}{\overline{\beta} + \omega} & \frac{\omega(1-p)}{\overline{\beta} + \omega} \end{bmatrix}$$
(7)

be the matrix denoting the transition probabilities from the transient states to the absorbing states. Then

$$(p_+, p_D, p_-) = (1, 0, 0)(Id - A)^{-1}R,$$
(8)

and so

$$\lambda_{\uparrow\uparrow} = \frac{p_+}{\tau} = \frac{\beta_b \omega p}{\beta_b + \overline{\beta} + \beta_f + \omega},\tag{9}$$

$$\mu = \frac{p_D}{\tau} = \frac{\overline{\beta} \left(\beta_b + \beta_f\right)}{\beta_b + \overline{\beta} + \beta_f + \omega},\tag{10}$$

$$\lambda_{1\downarrow} = \frac{p_{-}}{\tau} = \frac{\beta_f \omega \left(1 - p\right)}{\beta_b + \overline{\beta} + \beta_f + \omega}.$$
(11)



Figure 3: MT-dependent rates and probabilities for a system with $N_L = 8$ and $N_R = 6$ MTs growing from the left and right sides of maximum length N = 600, and polymerization, depolymerization, rescue and catastrophe rates equal to $\alpha = 2000$, $\beta = 5000$, $\alpha' = 5$, $\beta' = 15$.

Similarly, denote by ρ the reattachment rate for a motor in the diffusing state, which we assume from the *i*th position can only reattaching into the 11 state at the *i*th dimer. Further, assume N_a available MTs at the *i*th position. Treating the 11 state as absorbing, we have

$$\frac{1}{\rho} = \frac{2}{\omega} + \frac{1 + \overline{\beta}/\omega}{\gamma N_a}.$$
(12)

In the $\omega \gg \gamma$ regime, we have

$$\frac{1}{\rho} \approx \frac{1 + \overline{\beta}/\omega}{\gamma N_a},\tag{13}$$

from which the effective reattachment rate is given by

$$r = \frac{\gamma}{1 + \overline{\beta}/\omega}.$$
(14)

3 Walkoff and reattachment rates

Denote by p'_i and q'_i the probability a MT in equilibrium is *i* dimers long in the polymerizing state or depolymerizing state, respectively. Denote by $\pi_i = p'_i + q'_i$ the probability an MT is exactly *i* dimers long in either state, and by $\Pi_i = \sum_k^N \pi_k$ the probability a MT is at least *i* dimers long.

Assuming N_L MTs growing from the left and N_R MTs growing from the right, and a linear dependence of the reattachment rate to the number of nearby MTs, the rate of reattaching to (any) MT *i* dimers away from the left boundary is equal to

$$R_i = r(N_L \Pi_i + N_R \Pi_{N-i}), \tag{15}$$

with a probability p_i of reattaching to one of the N_L MTs given by

$$p_i = \frac{N_L \Pi_i}{N_L \Pi_i + N_R P_{N-i}},\tag{16}$$

and a probability $1 - p_i$ or reattaching to one of the N_R MTs. The rate of reattachment to a rightgrowing MT is therefore

$$r_i = p_i R_i,\tag{17}$$

and the rate of reattachment to a leftgrowing MT is $R_i(1-p_i)$.

For a motor attached i dimers along an MT, we approximate the walkoff probability f_i by the probability a sufficiently long MT is exactly i dimenselong,

$$f_i = \frac{\pi_i}{\Pi_i}.$$
(18)

When p'_i and q'_i are obtained from the equilibrium distribution of the dynamic MT model, the resulting reattachment rates and walkoff probabilities can significantly vary with position, as can be seen in Figure 3.

4 Absorption time probability density function

The absorption time of the motor is distributed according to a phase distribution [1] parametrized by (ζ, Q) , for which the probability density is given by

$$f(t) = \zeta^T e^{Qt} e \tag{19}$$

where ζ^T is a 3N-long row vector denoting the initial motor distribution, Q is the $3N \times 3N$ subgenerator matrix specifying the rates out of, and between, transient states of the chain, and e is the $3N \times 1$ vector denoting the rates from transient states to the absorbing state. For this distribution, the first and second moments are given by

$$\mathbb{E}\tau^T = -\zeta^T Q^{-1} \mathbf{1},\tag{20}$$

$$\mathbb{E}\tau^2 = 2\zeta Q^{-2}\mathbf{1}.\tag{21}$$

Specifically, for the first row in Figure 1, we have

$$Q_{i,i} = -(\lambda_{|\uparrow} + \lambda_{|\downarrow} + \mu), \qquad i \le N - 1, \qquad (22)$$

$$Q_{i,i+1} = (1 - f_i)\lambda_{\uparrow\uparrow}, \qquad i \le N - 2, \qquad (23)$$

$$Q_{i,i-1} = \lambda_{\uparrow\downarrow}, \qquad 0 < i < N - 1, \qquad (24)$$

$$Q_{i,i-1} - \lambda_{[1]}, \qquad \qquad 0 < i \le N-1, \qquad (24)$$
$$Q_{i,N+i} = f_i \lambda_{[1]} + \mu, \qquad \qquad i < N-1, \qquad (25)$$

$$e_0 = \lambda_{11}, \tag{26}$$

$$c_0 = \lambda |_{\downarrow}, \tag{20}$$

$$e_{N-1} = (1 - J_{N-1})\lambda_{\uparrow\uparrow}.$$
 (27)

By symmetry, for the third row in Figure 1, we have

$$Q_{2N+i,2N+j} = Q_{N-1-i,N-1-j}, \qquad 0 \le i < N, \qquad (28)$$
$$e_{2N+i} = e_{N-1-i}, \qquad 0 \le i \le N-1. \qquad (29)$$

$$e_{2N+i} - e_{N-1-i}, \qquad \qquad 0 \le i \le N-1.$$

Finally, for the middle row in Figure 1 (corresponding to the diffusing state), we have

$$Q_{N+i,i} = p_i R_i, \qquad 0 \le i \le N-1, \qquad (30)$$

$$Q_{N+i,2N+i} = (1-p_i)R_i, \qquad 0 \le i \le N-1, \qquad (31)$$

$$Q_{N+i,N+i} = -(R_i + 2d), \qquad 0 \le i \le N-1, \qquad (32)$$

$$Q_{N+i,N+i+1} = d, \qquad 0 \le i \le N-2, \qquad (33)$$

$$Q_{N+i,N+i-1} = d, \qquad 1 \le i \le N-1, \qquad (34)$$

$$e_N = e_{2N-1} = d, \qquad (35)$$

$$e_N = e_{2N-1} = d,$$



Figure 4: Left: The empirical absorption time distribution for a dynein motor simulated using the full model outlined in the main body of the paper, initialized in the middle of the MT bundle in a diffusing state. Away from the tail and near the bulk of the total probability, there is good agreement between the empirical distribution, and the one given in (19). Right: Using the formula in (19), we show the expended absorption time for a motor initialized i dimers away from the left boundary of the bundle, as well as its standard deviation (shaded region). As can be seen, the expected dynein absorption time is small and has small variance (compared to kinesin).



Figure 5: Trident-shaped matrix \mathcal{A} . Gray lines denote block matrix boundaries.

and $Q_{ij} = 0$ and $e_i = 0$ otherwise.

As can be seen in Figure 4, this simplified PDF allows us to approximate the true absorption time PDF, and investigate how the expected arrival time and its variances changes with the initial motor position.

5 Exact absorption time for a simplified system

Assume an equal number of MTs of max length N = 2k on each side, and no random detachment (i.e. the only way a motor can detach from a MT is by walking off). Denote by f_i that a step from the *i*th dimer on a right-growing MT results in a walk off, and denote by p_i the probability that a motor reattaching from the *i*th dimer from the left attaches to a right-growing MT. By symmetry, the submatrix of the generator corresponding to rates from transient states to transient states may be reduced from a $3N \times 3N$ matrix to a $1.5N \times 1.5N$ one, where only half the diffusing states are accounted for (the linear system may be rewritten in terms of the first N/2 diffusing states). Denote this pitchfork-shaped matrix (see Figure 5) by \mathcal{A} .



Figure 6: The case of the simplified MT model with fast reattachment and no random detachment can be seen as the simplest comparison of the walkoff effect. From (38) we have that the expected absorption time from the *i*th dimer grows as $1/\lambda$ times the doubled negative row sum of the *i*th row of some matrix X_N . We compare the maximal and minimal expected absorption time for a kinesin motor in a N-dimer long bundle with the maximal absorption time of a dynein motor. As can be seen from the graph above, a kinesin motor's expected absorption time is orders of magnitude greater than that of a dynein motor's. Furthermore, although the standard deviation of the kinesin absorption time varies considerably with the motor's location, it is generally very large.

The expected arrival times from the *i*th position are given by $\tau_i = -(\mathcal{A}\mathbf{2})_i$, where

$$\mathcal{A}(\lambda,r) = \begin{bmatrix} \begin{matrix} -\lambda & \lambda(1-f_0) & 0 & 0 & 0 & 0 \\ 0 & -\lambda & \lambda(1-f_1) & 0 & 0 & 0 & 0 \\ 0 & 0 & -\lambda & \lambda(1-f_2) & 0 & 0 & 0 & 0 & f_1\lambda & 0 \\ 0 & 0 & 0 & -\lambda & \lambda(1-f_3) & 0 & 0 & 0 & f_2\lambda \\ 0 & 0 & 0 & 0 & -\lambda & \lambda(1-f_4) & 0 & 0 & f_3\lambda & 0 \\ 0 & 0 & 0 & 0 & -\lambda & \lambda(1-f_4) & 0 & f_4\lambda & 0 \\ 0 & 0 & 0 & 0 & 0 & -\lambda & f_5\lambda & 0 & 0 \\ \hline rp_0 & 0 & 0 & 0 & rp_5 & -r(p_0+p_5) & 0 & 0 \\ 0 & rp_1 & 0 & 0 & rp_4 & 0 & 0 & -r(p_1+p_4) & 0 \\ 0 & 0 & rp_2 & rp_3 & 0 & 0 & 0 & 0 & -r(p_2+p_3) \\ \end{bmatrix}.$$
(36)

This matrix can be written as a 2×2 block matrix

$$\mathcal{A}(\lambda, r) = \begin{bmatrix} \lambda A & \lambda B \\ rC & rD \end{bmatrix},\tag{37}$$

which can be inverted by follows:

$$\begin{bmatrix} \lambda A & \lambda B \\ rC & rD \end{bmatrix}^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1}/\lambda & -(A - BD^{-1}C)^{-1}BD^{-1}/r \\ -D^{-1}C(A - BD^{-1}C)^{-1}/\lambda & (D^{-1} + D^{-1}C(A - D^{-1}C)^{-1}BD^{-1})/r \end{bmatrix}.$$
 (38)

In this case we have $(A(\lambda, r))^{-1} = \frac{1}{\lambda}X + \frac{1}{r}Y$, i.e. \mathcal{A} is decoupled from the system parameters, except for the ones governing the reattachment probabilities.

To investigate the walkoff effect, we don't need to consider the small effect of reattachment waiting times, so for convenience we assume $r \gg \lambda$, and set the block matrices in the right column to zero. In this case, the only nontrivial matrix we need to compute is $(A - BD^{-1}C)^{-1} = S^{-1}$. The $2k \times 2k$ matrix S has the shape given in Figure 7, and cannot be be inverted in a straightforward way. Instead we consider a decomposition into four $k \times k$ matrices:

$$S = \begin{bmatrix} S_1 & S_2 \\ S_3 & S_4 \end{bmatrix}.$$
(39)



Figure 7: Shape of $S = A - BD^{-1}C$.



Figure 8: Shape of $S_1 - S_2 S_4^{-1} S_3$.

This matrix can be similarly inverted blockwise:

$$S^{-1} = \begin{bmatrix} (S_1 - S_2 S_4^{-1} S_3)^{-1} & -(S_1 - S_2 S_4^{-1} S_3)^{-1} S_2 S_4^{-1} \\ -S_4^{-1} S_3 (S_1 - S_2 S_4^{-1} S_3)^{-1} & (S_4^{-1} + S_4^{-1} S_3 (S_1 - S_4^{-1} S_3)^{-1} S_2 S_4^{-1} \end{bmatrix}.$$
 (40)

Because \mathcal{A} is the transition rate matrix for the transient states, all of the elements in its inverse are negative. So it follows that for a lower bound on the absorption times from any of the first k right-moving states, we only need to compute

$$E^{-1}(-\mathbf{2}) = (S_1 - S_2 S_4^{-1} S_3)^{-1}(-\mathbf{2}).$$
(41)

The shape E is shown in Figure 8. Unfortunately it is not lower-triangular; otherwise we would be able to write down the solution recursively.

Fortunately, E is almost triangular. We divide the matrix into block matrices one final time, this time into a vector, lower-triangular matrix, constant, and row-vector:

$$E = \begin{bmatrix} x & L \\ c & y^T \end{bmatrix},\tag{42}$$

where

$$x_0 = -1 + f_0 p_0 - \frac{1 - f_0 f_{N-1} p_0 p_{N-1}}{1 - f_{N-1} p_{N-1}},$$
(43)

$$x_{i} = \frac{f_{i}f_{N-1}p_{0}p_{N-1-i}}{1 - f_{N-1}p_{N-1}} \prod_{j=N-1-i}^{N-2} \left(\frac{1 - f_{j}}{1 - f_{j}p_{j}}\right), \qquad 1 \le i \le k-2, \qquad (44)$$

$$y_{i} = \frac{f_{N-2-i}p_{i+1}(f_{k-1}p_{k} - f_{k-1} + 1)}{1 - f_{k}p_{k}} \prod_{j=k+1}^{N-3-i} \left(\frac{1 - f_{j}}{1 - f_{j+1}p_{j+1}}\right), \qquad 0 \le i \le k-2, \qquad (45)$$

$$y_{k-1} = -1 + f_{k-1}p_{k-1} - \frac{f_k p_{k-1}(f_k p_{k-1}(f_{k-1}p_k - f_{k-1} + 1))}{f_k p_k - 1},$$
(46)

$$c = \frac{f_{N-1}p_0(1 - f_{k-1} + f_{k-1}p_k)}{1 - f_{N-1}p_{N-1}} \prod_{j=k}^{N-2} \left(\frac{1 - f_j}{1 - f_j p_j}\right),\tag{47}$$

and L expanded to first order in small p and f entries is

$$L = \begin{bmatrix} 1 - p_0 & 0 & 0 & 0 & 0 \\ f_1 p_1 - 1 & 1 - f_1 & 0 & 0 & 0 \\ 0 & f_2 p_2 - 1 & 1 - f_2 & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & f_k p_k - 1 & 1 - f_{k-2} \end{bmatrix},$$
(48)

with the two main diagonals are exact, and the approximated by zero entries (i.e. the rest of the lower triangular part) are of order $f_i f_{N-i-1} p_i p_{N-i-1}$ in the exact matrix.

We can compute the inverse of E by flipping the columns, applying the standard inversion formula, and then switching the rows:

$$E^{-1} = \begin{bmatrix} -c^{-1}y^{T}U^{-1} & c^{-1} + c^{-1}y^{T}U^{-1}xc^{-1} \\ U^{-1} & -U^{-1}xc^{-1} \end{bmatrix},$$
(49)

where

$$U = L - \frac{1}{c} x y^T. ag{50}$$

Because every element of E^{-1} is nonpositive, a lower bound for the absorption time from the *i*th state for 1 < i < k may be obtained just from U^{-1} . By the Sherman-Morrison formula, U^{-1} may be viewed as a correction of L^{-1} :

$$U^{-1} = L^{-1} + \frac{L^{-1}xy^T L^{-1}}{c - y^T L^{-1}x}.$$
(51)

By induction on the rows and the back-substituion algorithm, it can be shown that $1 < (L^{-1})_{ij} < 1 + \epsilon$ for all $j \leq i$, and 0 otherwise, with (I need to go over this again)

$$\epsilon = \frac{1}{1 - f_0} \prod_{i=1}^{k} \frac{1 - f_{i+} p_{i+1}}{1 - f_i},\tag{52}$$

which is not very large because k = N/2 (i.e. the falloff probability in the middle of the bundle is small).

6 Simplified walkoff and reattachment rates

For a closed-form comparison of kinesin and dynein absorption times, we compare the absorption times for simplified values of R_i , p_i , and f_i .

In particular, assume a constant reattachment rate

$$R_i = r, (53)$$

and a rightgoing MT reattachment probability which decays linearly,

$$p_i = 1 - \frac{i+1}{N+1}.$$
(54)

This is qualitatively similar to the full model, as can be seen from Figure 3c. To approximate f_i , we note that for values of i not close to N, $f_i \ll 1$. On the other hand, for $i \leq N$ and

$$\frac{1+\alpha'/\beta}{1+\beta'/\alpha} \approx 1,\tag{55}$$

we have

$$f_i \approx \frac{1}{N+1-i}.\tag{56}$$

We therefore approximate f_i as in (56). The good agreement between this approximation and the falloff probabilities for an arbitrary choice of parameters can be seen in Figure 3).

With this choice of parameters, we have

$$f_i p_i = \frac{1}{N+1-i} \left(\frac{N-i}{N+1}\right) \approx \frac{1}{N+1}$$
(57)

for most $i \ll N$, and

$$\prod_{i=k}^{N-2} (1-f_i) = \prod_{i=k}^{N-2} \frac{N-i}{N+1-i} = \frac{4}{N+2}.$$
(58)

From this we have

$$c \approx \frac{1}{2(k+1)} \left(1 + \frac{1}{2k}\right)^k \lesssim \frac{e^{1/2}}{2(k+1)},$$
(59)

where the additional factor of 1/2 as a correction to (57) in the large *i* case. Using the exact expressions for *x* and *y*, we can show that $y^T L^{-1}x$ is negligibly small compared to *c*. Back-substituting this into the matrices U^{-1} and E^{-1} (which can simplified using the observation in Section 8), we see that the expected absorption time for kinesin must grow at least as quickly as N^3 , i.e. much faster than the growth of dynein expected absorption time.

In the case that these simplified rates are not used, we have that the probability of reattachment to a right-going MT near the right boundary is even smaller, and the rate of reattachment is higher; thus we expected the walk-off effect to be even more pronounced. Indeed, numerical simulations suggest the kinesin expected absorption time grows exponentially.

7 Outer product notes

To approximate U^{-1} , we first analyze x and y, in order to estimate the matrix xy^T . For the vector x we have the property that

$$\frac{x_0}{x_1} = \frac{(f_{N-2}p_{N-2}-1)(f_{N-1}p_{N-1}-1)\left(-\frac{f_0f_{N-1}p_0p_{N-1}}{f_{N-1}p_{N-1}-1}+f_0p_0-1\right)}{f_1f_{N-1}p_0p_{N-2}(1-f_{N-2})},\tag{60}$$

$$\frac{x_i}{x_{i+1}} = \frac{f_i p_{N-1-i} (1 - f_{N-2-i} p_{N-2-i})}{f_{i+1} p_{N-2-i} (1 - f_{N-2-i})}, \quad i > 0.$$
(61)

In particular, for falloff and reattachment probabilities which do not increase/decrease dramatically between two adjacent dimers, we have that $|x_0| \gg |x_i|$ for any i > 0, because $0 < f_1, p_{N-2} \ll 1$. Similarly, for the vector y we have

$$\frac{y_{k-2}}{y_{k-3}} = \frac{(f_k p_k - 1) \left(f_{k+1} p_{k+1} - 1\right) \left(f_{k-1} p_{k-1} - \frac{f_k p_{k-1} (f_{k-1} p_k - f_{k-1} + 1)}{f_k p_{k-1}} - 1\right)}{f_{k+1} p_{k-2} \left(1 - f_k\right) \left(f_{k-1} p_k - f_{k-1} + 1\right)},\tag{62}$$

$$\frac{y_{i+1}}{y_i} = \frac{f_{N-3-i}p_{i+2}(1 - f_{N-2-i}p_{N-2-i})}{f_{N-2-i}p_{i+1}(1 - f_{N-2-i-1})}, \quad i < k-3,$$
(63)

from which it follows that $|y_{k-2}| \gg y_i$ for all i < k-2, assuming $f_{k+1} \ll 1$ (this is the walkoff probability in the middle). We can therefore very roughly approximate the outer product by a zero matrix with a non-zero upper right corner, equal to x_0y_{k-2} .

References

[1] Marcel F Neuts. Matrix-geometric solutions in stochastic models: an algorithmic approach. Courier Corporation, 1994.