Supplemental Information - Contributions of cell growth and biochemical reactions to non-genetic variability of cells

Anne Schwabe & Frank J. Bruggeman

Systems Bioinformatics, VU University, De Boelelaan 1087, 1081 HV, Amsterdam, The Netherlands

1 Decomposition of the variance in molecule copy numbers

As a first step we use the law of total variance to split the variance into a cell cycle stage dependent and independent part:

$$\langle\delta^2 x\rangle = \langle\delta^2\langle x|a\rangle\rangle + \langle\langle\delta^2 x|a\rangle\rangle = \underbrace{\langle\delta^2\langle x_a\rangle\rangle}_{\substack{\text{variance due to}\\ \text{differences in}\\ \text{cell cycle stage}}} + \langle\langle\delta^2 x_a\rangle\rangle \tag{S-1}$$

The variance at cell cycle stage a can further be split into the variance deriving from the molecules that the cell inherited from its mother and that are not yet degraded, denoted $\bar{x}_0$, the variance that stems from newly synthesized molecules, and their covariance:

$$\langle\delta^2 x_a\rangle = \langle\delta^2\bar{x}_0\rangle + \underbrace{\langle\delta^2 X_a\rangle}_{\substack{\text{variance in synthesis and extrin-}\\ \text{sic noise in newly synthesized}\\ \text{molecules (”networking”)}}} + 2cov(\bar{x}_0, X_a) \tag{S-2}$$

where the covariance term becomes non-zero if the synthesis rate depends on “extrinsic” factors. The first term in eq. S-2 can also be decomposed using the law of total variance: for a given number of molecules at cell birth the variance at cell cycle stage a equals $p(a)(1-p(a))\langle x_0\rangle$, which is the variance of a binomial distribution. This is because each molecule has an independent survival probability, giving rise to a binomial distribution of remaining molecules for a given number of molecules to start with. Since x_0 itself is distributed, the variance of the average at cell cycle stage a must be added to this:

$$\langle\delta^2\bar{x}_0\rangle = p(a)(1-p(a))\langle x_0\rangle + p(a)^2\langle\delta^2 x_0\rangle \tag{S-3}$$

From the law of total variance we obtain for the variance at cell birth:

$$\langle \delta^2 x_0 \rangle = \underbrace{\langle \langle \delta^2 x_0 | x_T \rangle \rangle}_{\text{partitioning variance}} + \underbrace{\langle \delta^2 \langle x_0 | x_T \rangle \rangle}_{\text{variance due to mother cell variability}} \tag{S-4}$$

Combining these equations and taking the appropriate averages yields Eq. 6 in the main text:

$$\begin{aligned}
\langle \delta^2 x \rangle &= \overbrace{\overbrace{\underbrace{\langle p(a)^2 \rangle \langle \langle \delta^2 x_0 | x_T \rangle \rangle}_{\text{partitioning variance}} + \underbrace{\langle p(a)^2 \rangle \langle \delta^2 \langle x_0 | x_T \rangle \rangle}_{\text{variance due to mother cell variability}}}^{\text{Decayed variance of molecules at cell birth, } \langle p(a)^2 \rangle \langle \delta^2 x_0 \rangle} + \underbrace{\langle \delta^2 \langle x_a \rangle \rangle}_{\text{variance due to differences in cell cycle stage}}}^{\text{Growth-induced variance}} \\
&+ \underbrace{\underbrace{\langle p(a)(1-p(a)) \rangle \langle x_0 \rangle}_{\text{variance in the degradation of molecules obtained at division}} + \underbrace{\langle \langle \delta^2 X_a \rangle \rangle}_{\text{variance in synthesis and extrinsic noise in newly synthesized molecules (“networking”)}} + \underbrace{2 \langle p_a cov(x_0, X_a) \rangle}_{\text{covariance due to dependence on extrinsic factors}}}_{\text{Reaction-induced variance}}
\end{aligned} \tag{S-5}$$

where we used $cov(\bar{x_0}, X_a) = p(a) cov(x_0, X_a)$.With $p(a) = e^{-k_d a}$, the averages $\langle p(a)^2 \rangle$ and $\langle p(a)(1-p(a) \rangle$ can be determined from the cell cycle stage distribution (eq. S-35). For deterministic interdivision times we obtain:

$$\begin{aligned}
\langle p(a)^2 \rangle &= \int_a u(a) e^{-2 k_d a} da = \frac{\left(2 - 4^{-\frac{k_d}{\mu}}\right)}{2 k_d/\mu + 1} \\
\langle p(a)(1-p(a)) \rangle &= \int_a u(a) e^{-k_d a} \left(1 - e^{-k_d a}\right) da \\
&= \left(\frac{2 - 2^{-\frac{k_d}{\mu}}}{k_d/\mu + 1} + \frac{-2 + 4^{-\frac{k_d}{\mu}}}{2 k_d/\mu + 1} \right)
\end{aligned} \tag{S-6}$$

2 Variance from cell division and partitioning

To calculate the variance at cell birth when the partition ratio, q, is not deterministic but follows a distribution with mean $\langle q \rangle = 1/2$ and variance $\langle \delta^2 q \rangle$, the law of total variance is applied twice under the assumption that x_T and q are independent:

$$\begin{aligned}
\langle \delta^2 x_0 \rangle &= \langle \langle \delta^2 x_0 | q \rangle \rangle + \langle \delta^2 \langle x_0 | q \rangle \rangle && \text{(S-7)} \\
\langle \delta^2 x_0 | q \rangle &= \langle \langle \delta^2 x_0 | x_T \rangle \rangle + \langle \delta^2 \langle x_0 | x_T \rangle \rangle && \text{(S-8)} \\
&= \langle q(1-q) x_T \rangle + \langle \delta^2 \langle q x_T \rangle \rangle \\
&= q(1-q)\langle x_T \rangle + q^2 \langle \delta^2 x_T \rangle
\end{aligned}$$

Combining equations S-7, S-8, and $\langle q \rangle = 1/2$ yields:

$$\begin{aligned}
\langle \delta^2 x_0 \rangle &= \langle q(1-q)\langle x_T \rangle \rangle + \langle q^2 \langle \delta^2 x_T \rangle \rangle + \langle \delta^2 (q \langle x_T \rangle) \rangle && \text{(S-9)} \\
&= \langle q(1-q) \rangle \langle x_T \rangle + \langle q^2 \rangle \langle \delta^2 x_T \rangle + \langle \delta^2 q \rangle \langle x_T \rangle^2 \\
&= \langle \delta^2 q \rangle \langle x_T \rangle^2 + \left(\frac{1}{4} - \langle \delta^2 q \rangle \right) \langle x_T \rangle + \left(\frac{1}{4} + \langle \delta^2 q \rangle \right) \langle \delta^2 x_T \rangle
\end{aligned}$$

3 Calculation of total variance from the partitioning mechanism and the variance in the synthesis process

The mechanism of partitioning of molecules at division, together with the condition of balanced growth, sets the boundary conditions for the distributions (and their moments) of copy numbers at birth and at division. Under the assumption of deterministic interdivision times (relaxation of this assumption is discussed in the main text), these boundary conditions can be combined with the kinetics of molecule turnover to yield a system of equations for the moments of the copy number distribution at cell cycle stage a:

By setting $a = T$ in eq. S-2 and combining with eqs. S-4, S-5 one can express the population level variance solely in terms of the variance in the synthesis, $\langle \delta^2 X_a \rangle$ (which can include intrinsic and extrinsic components) and its correlation with the number of molecules at cell birth, the variance in the partition distribution, $\langle \delta^2 q \rangle$, and the survival probability of molecules:

$$\langle x_a \rangle = \underbrace{\frac{\langle X|T \rangle \langle q \rangle}{1 - p(T)\langle q \rangle}}_{\langle x_0 \rangle} p(a) + \langle X|a \rangle \qquad \text{(S-10)}$$

$$\langle \delta^2 x_a \rangle = p(a)^2 \langle \delta^2 x_0 \rangle + p(a)(1-p(a)) \frac{\langle X|T \rangle \langle q \rangle}{1 - p(T)\langle q \rangle} + \langle \delta^2 X_a \rangle + 2p(a) cov(x_0 | X_a) \qquad \text{(S-11)}$$

$$\langle \delta^2 x_0 \rangle = \frac{\langle x_T \rangle \langle q \rangle - \langle x_T \rangle \langle q^2 \rangle + \langle \delta^2 q \rangle \langle x_T \rangle^2}{1 - \langle q^2 \rangle p(T)^2} + \frac{\langle q^2 \rangle (p(T) - p(T)^2) \langle x_0 \rangle + \langle q^2 \rangle \langle \delta^2 X_T \rangle + 2p(T) cov(x_0 | X_T)}{1 - \langle q^2 \rangle p(T)^2}$$

Marginalizing out the cell cycle stage contribution the total variance, $\langle \delta^2 x \rangle$, can be calculated from these equations. While the variance decomposition is meant to be used on time-resolved data, the

above equations allow for an estimation of the relative magnitude of the different variance contributions in a sample of extant cells (as would be observed in a typical FACS or FISH experiment).

3.1 Contribution of partitioning variance for a stable molecule

Here we calculate the variance for a stable molecule ($k_{deg} = 0$) that is synthesized by a zero order reaction (with rate constant k_s), where the partitioning ratio at cell division, q, fluctuates. The average and variance at cell birth are given by

$$\langle x_0 \rangle = \langle q \rangle \langle x_T \rangle = \langle q \rangle (\langle x_0 \rangle + \langle X|T \rangle) = \langle X|T \rangle \tag{S-12}$$

$$\langle \delta^2 x_0 \rangle = \langle \delta^2 q \rangle \langle x_T \rangle^2 + (1/4 - \langle \delta^2 q \rangle)\langle x_T \rangle + (1/4 + \langle \delta^2 q \rangle) \underbrace{\langle \delta^2 x_T \rangle}_{= \langle \delta^2 x_0 \rangle + \langle \delta^2 X|T \rangle} \tag{S-13}$$

With $\langle x_T \rangle = 2\langle x_0 \rangle = 2k_sT$ and $\langle \delta^2 X|T \rangle = k_sT$ this yields:

$$\langle \delta^2 x_0 \rangle = \frac{k_sT(-3 + (4 - 16k_sT)\langle \delta^2 q \rangle)}{-3 + 4\langle \delta^2 q \rangle} \tag{S-14}$$

Combining this with equations S-1 and S-2 yields for the population level variance:

$$\langle \delta^2 x \rangle = \langle x \rangle + \langle x \rangle^2 (1 - 2\ln(2)^2) + \langle x \rangle^2 \frac{16\ln(2)^2 \langle \delta^2 q \rangle)}{3 - 4\langle \delta^2 q \rangle} \tag{S-15}$$

4 Contribution of the cell cycle stage distribution to variance in copy numbers

If the synthesis and degradation rates are constant throughout the cell cycle, the cell cycle stage dependent average of net production is given by

$$\langle X_a \rangle = \frac{k_s}{k_d}(1 - e^{-k_d a}) \tag{S-16}$$

With $\langle x_a \rangle = p(a)\langle x_0 \rangle + \langle X_a \rangle$ and $\langle x_0 \rangle = 1/2\langle x_T \rangle$ the average number of molecules at cell birth equals:

$$\langle x_0 \rangle = \frac{\frac{k_s}{2k_d}(1 - e^{-k_d T})}{1 - e^{-k_d T}/2} \tag{S-17}$$

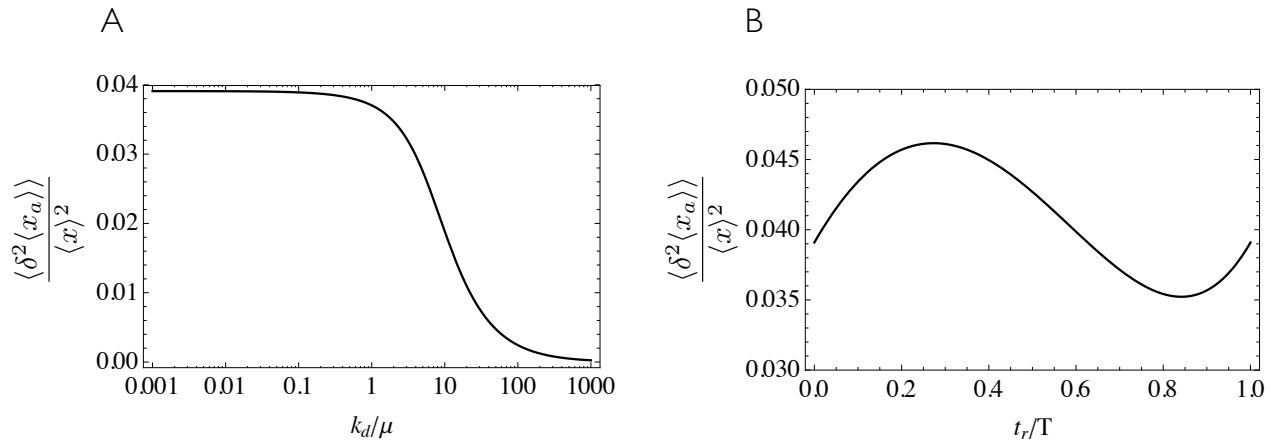

Figure S-1: **Noise contribution of differences in average copy number throughout the cell cycle** A) constant synthesis rate throughout the cell cycle. B) synthesis rate doubles at time t_r ($k_d = 0$)

From this, the cell cycle stage contribution to variance can be calculated as (Fig. S-1 A):

$$\begin{aligned}
\langle \delta^2 \langle x_a \rangle \rangle &= \int_0^T u(a)(\langle x_a \rangle - \langle x \rangle)^2 da \\
&= \int_0^T u(a) \left(e^{-k_d a} \frac{\frac{k_s}{2k_d}\left(1 - e^{-k_d T}\right)}{1 - e^{-k_d T}/2} + \frac{k_s}{k_d}\left(1 - e^{-k_d a}\right) - \frac{k_s}{k_d + \frac{\log(2)}{T}} \right)^2 da \\
&= \langle x \rangle^2 \frac{\left(2^{2k_d/\mu+1} - 1\right)(k_d/\mu)^2 - 4k_d/\mu \left(2^{k_d/\mu} - 1\right)^2 - 2\left(2^{k_d/\mu} - 1\right)^2}{\left(2^{k_d/\mu+1} - 1\right)^2 (k_d/\mu)^2 (2(k_d/\mu) + 1)}
\end{aligned} \tag{S-18}$$

with $\mu = ln(2)/T$ and $\langle x \rangle = \frac{k_s}{\mu + k_d}$.

When the synthesis rate doubles at some point during the cell cycle due to replication of the gene, the cell cycle stage dependent variance for a stable molecule can be expressed as (Fig. S-1 B):

$$\begin{aligned}
\langle x_a \rangle &= k_s(a + 2T - t_r) \quad \text{for } 0 < a < t_r \\
\langle x_a \rangle &= 2k_s(a + T - t_r) \quad \text{for } t \leq a < T \\
\langle \delta^2 \langle x_a \rangle \rangle &= \langle x \rangle^2 \left(-1 + 2^{\frac{t_r}{T}}(3 + \ln(4)) \right) \\
&\quad - \langle x \rangle^2 \frac{4^{\frac{t_r}{T}}}{2} \left(\left(\frac{t_r}{T} \right)^2 \ln(2)^2 - 2\frac{t_r}{T} \ln(2)(1 + \ln(4)) \right) \\
&\quad - \langle x \rangle^2 \frac{4^{\frac{t_r}{T}}}{2} \left(2 + 4\ln(2)^2 + \ln(16) \right)
\end{aligned} \tag{S-19}$$

5 Population level variance decomposition for a simple burst model

Here we apply the results from the previous sections to a burst model where the times between bursts are exponentially distributed (with average time between bursts equal to $1/k_s$), burst sizes with a general distribution (average burst size $\langle b\rangle$ and variance $\langle\delta^2 b\rangle$). For simplicity we take the molecule to be stable, i.e. $k_d = 0$. The average and variance of newly made molecules at cell cycle stage a for this model are given by [1]:

$$\langle X_a\rangle = k_s a\langle b\rangle \tag{S-20}$$
$$\langle\delta^2 X_a\rangle = k_s a(\langle\delta^2 b\rangle + \langle b\rangle^2) \tag{S-21}$$

First we solve for the mean and variance at cell birth:

$$\langle x_0\rangle = \langle q\rangle\langle x_T\rangle = \langle q\rangle(\langle x_0\rangle + k_s T\langle b\rangle) = \frac{\langle q\rangle k_s T\langle b\rangle}{1-\langle q\rangle} = k_s T\langle b\rangle \tag{S-22}$$

$$\langle\delta^2 x_0\rangle = \overbrace{\langle\delta^2 q\rangle\langle x_T\rangle^2}^{\langle\delta^2\langle x_0|q\rangle\rangle,\text{ variance due to } q \text{ fluctuations}} + \overbrace{\left(\frac{1}{4} - \langle\delta^2 q\rangle\right)\langle x_T\rangle}^{\text{partitioning variance}} + \overbrace{\left(\frac{1}{4} + \langle\delta^2 q\rangle\right)\langle\delta^2 x_T\rangle}^{\text{variance due to mother cell heterogeneity}} =$$

$$\frac{\langle\delta^2 q\rangle(2k_s T\langle b\rangle)^2 + \left(\frac{1}{4} - \langle\delta^2 q\rangle\right)(2k_s T\langle b\rangle) + \left(\frac{1}{4} + \langle\delta^2 q\rangle\right)\langle\delta^2 X_T\rangle}{1 - \frac{1}{4} - \langle\delta^2 q\rangle} \tag{S-23}$$

The variance due to the cell cycle stage distribution is given by

$$\begin{aligned}\langle\delta^2\langle x_a\rangle\rangle &= \int u(a)(\langle x_a\rangle - \langle x\rangle)^2 da \\ &= \int u(a)\left(x_0 + k_s a\langle b\rangle - \frac{k_s T\langle b\rangle}{ln(2)}\right)^2 da \\ &= \langle x\rangle^2(1 - 2ln(2)^2)\end{aligned} \tag{S-24}$$

and the average variance due to new synthesis equals:

$$\begin{aligned}\langle\langle\delta^2 X_a\rangle\rangle &= \int u(a)k_s a(\langle\delta^2 b\rangle + \langle b\rangle^2)da \\ &= \left(\frac{1}{\ln 2} - 1\right)k_s T\left(\langle\delta^2 b\rangle + \langle b\rangle^2\right)\end{aligned} \tag{S-25}$$

Combining these equations with S-1 yields:

$$\langle \delta^2 x \rangle = \overbrace{\langle x \rangle^2 (1 - 2ln(2)^2)}^{\text{variance due to cell cycle stage distribution}} + \overbrace{\left(\frac{1}{\ln 2} - 1\right) k_s T \left(\langle \delta^2 b \rangle + \langle b \rangle^2\right)}^{\text{variance due to synthesis}} + \underbrace{\frac{k_s T(\langle b \rangle(2 + \langle b \rangle) + \langle \delta^2 b \rangle + 4(\langle b \rangle(-2 + \langle b \rangle + 4\langle b \rangle k_s T) + \langle \delta^2 b \rangle)\langle \delta^2 q \rangle)}{3 - 4\langle \delta^2 q \rangle}}_{\langle \delta^2 x_0 \rangle \text{ variance at cell birth}}$$

6 Variance decomposition with extrinsic noise

We model extrinsic noise as temporal fluctuations in the synthesis rate, $k_s(t)$, of molecule $\mathbb{X}$. Considering only the net new synthesis, the average and variance at cell cycle state a is given by [2]:

$$\langle X_a \rangle = \int_0^a p(a-t)\langle k_s(t) \rangle \tag{S-26}$$

$$\langle \delta^2 X_a \rangle = \int_0^a p(a-t)(1 - p(a-t))\langle k_s(t) \rangle dt + \int_0^a \int_0^a p(a-t_1)p(a-t_2) cov(k_s(t_1), k_s(t_2)) dt_1 dt_2 \tag{S-27}$$

When modeling synthesis as a modulated Poisson process

$$\varnothing \overset{k_s Y(t)}{\rightarrow} X \overset{k_d}{\rightarrow} \varnothing \tag{S-28}$$

where the synthesis rate depends on a fluctuating species, $\mathbb{Y}$, the covariance function equals the sum of a dirac delta function accounting for the shot noise and a term accounting for the fluctuations in $Y(t)$:

$$cov(k_s Y(t_1), k_s Y(t_1 + \tau)) = k_s \langle Y(t_1) \rangle \delta(\tau) + k_s^2 cov(Y(t_1), Y(t_2)) \tag{S-29}$$

The variance in the number of newly synthesized molecules then equals:

$$\langle \delta^2 X_a \rangle = k_s \int_0^a p(a-t)\langle Y(t) \rangle dt + k_s^2 \int_0^a \int_0^a p(a-t_1)p(a-t_2) cov(Y(t_1), Y(t_2)) dt_1 dt_2 \tag{S-30}$$

The total copy number variance at cell cycle stage a is given by

$$\langle \delta^2 x_a \rangle = \langle \delta^2 x_0 \rangle + \langle \delta^2 X_a \rangle + 2cov(p(a)x_0, X_a) \tag{S-31}$$

where the covariance term accounts for fluctuations in the synthesis rate that span more than one generation. If there is no variance in the partition function, i.e. $q = 1/2$, the variance at cell cycle stage a can also be calculated as:

$$\begin{aligned}\langle \delta^2 x_a \rangle &= k_s \int_{-\infty}^{a} p(a-t) \left(\frac{1}{2}\right)^i \langle Y(t) \rangle dt + \\ & \quad k_s^2 \int_{-\infty}^{a} \int_{-\infty}^{a} p(a-t_1) p(a-t_2) \left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^k cov(Y(t_1), Y(t_2)) dt_1 dt_2 \end{aligned} \tag{S-32}$$

with $i = \lceil \frac{-t}{T} \rceil$, $j = \lceil \frac{-t_1}{T} \rceil$, $k = \lceil \frac{-t_2}{T} \rceil$, where $\lceil .. \rceil$ denotes the ceiling function. $t = 0$ defines the time point of the last cell division. For the system defined in Eq. S-28 assuming that Y varies with an exponentially decaying autocorrelation function with rate constant k_Y and $\langle \delta^2 Y \rangle = \langle Y \rangle$ the covariance equals:

$$cov(x_0, X_a) = \frac{k_s^2 \langle Y \rangle e^{a(-(k_Y + k_d))} \left(e^{k_Y a} - e^{k_d t}\right) \left(e^{T(k_Y + k_d)} - 1\right)}{(k_Y^2 - k_d^2)\left(2e^{T(k_Y + k_d)} - 1\right)} \tag{S-33}$$

7 Using generating functions to calculate the cell cycle stage dependent copy number distribution

7.1 Associated cell cycle stage distribution for a discrete interdivision time distribution

We consider a discrete interdivision time distribution, $f(t)$, which models division after deterministic time intervals, T:

$$f(t) = \delta(T - t) \tag{S-34}$$

with δ as the Dirac delta distribution such that $f(t = T) = 1$ and $f(t \neq T) = 0$. The cell cycle stage distribution $u(a)$ of an exponential growing population of cells with specific growth rate μ is related to the interdivision time distribution by (Eq. 8 in reference [3]),

$$\begin{aligned} u(a) &= 2 \cdot \mu \cdot e^{-\mu \cdot a} \int_a^{\infty} f(t) dt \\ &= \frac{1}{T} 2^{1 - \frac{a}{T}} \ln 2 \qquad \text{for: } 0 \leq a \leq T \end{aligned} \tag{S-35}$$

With $\mu = \frac{\ln 2}{T}$.

Note that we have the following relationships,

$$\begin{aligned}
\int_0^T u(a)da &= 1 \\
\langle a \rangle &= \int_0^T a \cdot u(a)da = T\left(\frac{1}{\ln 2} - 1\right) \approx 0.44T \\
\langle \delta^2 a \rangle &= \int_0^T a^2 \cdot u(a)da - \langle a \rangle^2 = T^2\left(\frac{1}{(\ln 2)^2} - 2\right) \\
\frac{\langle \delta^2 a \rangle}{\langle a \rangle^2} &= \frac{1 - 2(\ln 2)^2}{(\ln 2 - 1)^2} = 0.41
\end{aligned} \tag{S-36}$$

7.2 Background on generating functions

7.2.1 The probability generating function

Let X be a discrete random variable that takes non-negative integer values $X \in \{0, 1, 2, ..\}$. The (point) probability that X takes value i is defined as $p_i = P(X = i)$. The (probability) generating function of X denoted by $G_X(z)$ is defined as,

$$G_X(z) = \sum_{i=0}^{\infty} p_i z^i = \langle z^X \rangle \tag{S-37}$$

The moments can be obtained from,

$$\langle X^i \rangle = \frac{1}{z!}\left(\frac{d}{dz}\right)^i G_X(z)\Big|_{z=1} \tag{S-38}$$

7.2.2 The generating function of a sum of independent random variables, X and Y

The generating function of a sum of independent random variables, X and Y, equals

$$\begin{aligned}
G_{X+Y}(z) &= \langle z^{X+Y} \rangle = \langle z^X z^Y \rangle = \underbrace{\langle \delta z^X \delta z^Y \rangle}_{=0;\ \text{independence}} + \langle z^X \rangle \langle z^Y \rangle \\
&= \langle z^X \rangle \langle z^Y \rangle = G_X(z) G_Y(z)
\end{aligned} \tag{S-39}$$

It is said that the probability mass function (pmf) $P_Y(Y)$ is obtained from the convolution of the pmfs of X and Y. Convolution thus means the sum of random variables and can be obtained from the product of the generating functions.

7.2.3 A compound distribution and its generating function

Let Y be a sum of independent and identically distributed random variables X_i then

$$Y = X_1 + X_2 + ...X_N \tag{S-40}$$

and let N be a positive integer-valued random variable. We define $G_X(z)$ and $G_N(z)$ and $G_Y(z)$ now equals,

$$\begin{aligned} G_Y(z) &= \langle z^Y \rangle = \langle \langle z^Y | N \rangle \rangle = \langle \langle z^{X_1+X_2+...X_n} | N \rangle \rangle \\ &= \langle \langle z^{X_1} z^{X_2} ... z^{X_N} | N \rangle \rangle \\ &= \langle G_X(z)^N \rangle \\ &= G_N(G_X(z)) \end{aligned} \tag{S-41}$$

7.2.4 Bernoulli distribution and its generating function

A random variable X is Bernoulli distributed if,

$$X = \begin{cases} 1 & \text{when the event is successful, probability } p, \\ 0 & \text{when the event is successful, probability } q = 1 - p \end{cases} \tag{S-42}$$

The generating function is then $G_X z = p_0 z^0 + p_1 z^1 = q + pz = 1 - p + pz$.

We are going to need this generating function to calculate the probability for the number of molecules obtained at cell division and to calculate the probability for the number of molecules that have not been degraded within a certain time interval.

7.3 The generating function for the number of molecules in a population of cells engaged in balanced growth

Denote by $G_{(x,a)}(z)$ the probability generating function (pgf) for the number of molecules, x, at cell cycle stage a. $H_{(x_0,a)}(z)$ is the pgf for the distribution of molecules obtained at cell birth, x_0, that have survived until cell cycle stage a. $F_{(x,a)}(z)$ is the pgf of the number of molecules newly produced that have not yet been degraded during time a. Here we assume that fluctuations in the synthesis rate are limited to a single generation so that $F_{(x,a)}(z)$ and $H_{(x_0,a)}(z)$ are independent.

At any cell cycle stage a the number of molecules equals the sum of the number of molecules obtained from birth that have not yet been degraded and those have been newly produced and not yet been degraded. Thus, the pgf of the number of molecules at cell cycle stage a can be obtained from the convolution,

$$G_{(x,a)}(z) = H_{(x_0,a)}(z) F_{(X,a)}(z) \tag{S-43}$$

$H_{(x_0,a)}(z)$ is a compound generating function for the probability of the number of molecules that have not been degraded until cell cycle stage a of the x_0 molecules obtained at birth. With first order independent degradation of molecules the survival probability for each molecule equals $p(t) = e^{-k_d t}$. From this and equations S-41 and S-42, $H_{(x_0,a)}(z)$ can be expressed as

$$H_{(x_0,a)}(z) = G_{(x,0)}(1 - p(a) + p(a)z) \tag{S-44}$$

We assume independent (binomial) partitioning at division where one daughter cell receives molecules with probability q (the other daughter with $1 - q$). We consider that q follows a probability density function $g(q)$. The "partitioning probability" q can be influenced by the volume ratio of the two daughter cells or binding of molecules to intracellular compartments. The copy number probabilities at birth (a=0) in a daughter cell is related to the copy number probability at the time of division, T, in the mother cell,

$$H_{(x_0,0)}(z) = G_{(x,0)}(z) = \langle G_{(x,T)} \underbrace{(1 - q + qz)}_{\substack{\text{binomial partitioning} \\ \text{of the x molecules at} \\ \text{time T in the mother}}} \rangle \tag{S-45}$$

Combining equations S-43 to S-45 yields

$$\begin{aligned} H_{(x_0,a)}(z) &= G_{(x,0)}(1 - p(a) + p(a)z) \\ &= \langle G_{(x,T)}(1 - q + q(1 - p(a) + p(a)z)) \rangle \\ &= \langle G_{(x,T)}(1 - qp(a)(1 - z)) \rangle \qquad \text{(S-46)} \\ G_{(x,a)}(z) &= \langle G_{(x,T)}(1 - qp(a)(1 - z)) \rangle F_{(X,a)}(z) \qquad \text{(S-47)} \end{aligned}$$

If $q = 1/2$ (i.e. $g(q = 1/2) = 1$) eq. S-47 can be solved by iteration for $G_{(x,T)}$ (if q follows a distribution, the moments of the copy number distribution can still be obtained from eq. S-47):

$$\begin{aligned} G_{(x,T)}(z) &= G_{(x,T)}(1 - \frac{1}{2}p(T)(1 - z))F_{(X,T)}(z) \qquad \text{(S-48)} \\ &= G_{(x,T)}(1 - (\frac{1}{2}p(T))^2(1 - z))) \times F_{(X,T)}(1 - \frac{1}{2}p(T)(1 - z))F_{(X,T)}(z) \\ &= G_{(x,T)}(1 - (\frac{1}{2}p(T))^n(1 - z))) \times \prod_{i=0}^{n} F_{(X,T)}(1 - (\frac{1}{2}p(T))^i(1 - z)) \qquad \text{(S-49)} \end{aligned}$$

In the limit of $n \to \infty$ the first term in equation S-49 becomes one resulting in:

$$G_{(x,T)}(z) = \prod_{i=0}^{\infty} F_{(X,T)}(1 - (\frac{1}{2}p(T))^i(1 - z)) \tag{S-50}$$

Combining this with eq. S-47 and $F_{(X,0)}(z) = 1$ yields:

$$G_{(x,0)}(z) = \prod_{i=0}^{\infty} F_{(X,T)}(1 - \frac{1}{2}(\frac{1}{2}p(T))^i(1-z)) \quad \text{(S-51)}$$

$$G_{(x,a)}(z) = F_{(X,a)}(z) \times \prod_{i=0}^{\infty} F_{(X,T)}(1 - \frac{1}{2}p(a)(\frac{1}{2}p(T))^i(1-z)) \quad \text{(S-52)}$$

8 Simulation of cell populations in balanced growth

Reactions were simulated using the next reaction method with extension to volume dependent reactions [4]. The time of cell division (calculated from the volume at cell birth, the volume at division, and the cell's rate of volume increase) was added to the list of reaction times. Upon division, all molecules were distributed binomially over the two daughter cells with a probability equal to the ratio of daughter to mother volume. The complete lineage tree was simulated, either starting with a single cell (cell divisions with a distributed division fraction and with distributed rate of volume increase) or with a collection of cells where the remaining life length was calculated from the theoretically calculated interdivision time distribution according to [3] (cell divisions with distribution of volumes at division). The reason for this is that the latter mechanism on its own doesn't lead to a time invariant cell cycle stage distribution when starting with a single cell. Simulations were run until all distributions (cell cycle stage, copy number, volume) became stationary.

References

[1] L. Liu, B. R. K. Kashyap, and G. J. C. Templeton. On the GIX/G/∞ system. *Journal of Applied Probability*, 27:671–683, Sep 1990.

[2] I. Eliazar. On the discrete-time g/gi/∞ queue. *Probab. Eng. Inf. Sci.*, 22:557–585, October 2008.

[3] P. R. Painter and A. G. Marr. Mathematics of microbial populations. *Annu. Rev. Microbiol.*, 22:519–548, 1968.

[4] M.A. Gibson and J. Bruck. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J. Phys. Chem. A*, 2000.