1. THE MASTER EQUATION APPROACH

The master equation corresponds to the statement that the probability of being in a given state changes depending on the probabilities of transition to and from any other state in the system. It provides the full probability distribution when it can be directly solved. Unfortunately, this is not often the case, so we must settle for some of the moments of the distribution. These are easily obtained from the generating function, so we will work with the master equation in a form in which it depends on the generating function rather than the distribution.

The genetic network is defined by N state variables n_1 .. n_N and M rate constants k_1 .. k_M. The variables denote the number of copies of a certain chemical species such as mRNAs or proteins. Before applying the master equation approach to determine the noise properties of a genetic network we will start by obtaining the master equation in the generating function form for some elementary chemical equations:

Synthesis from a template

In numerous genetic reactions, such as transcription and translation, mRNAs and proteins are synthesized from a template (DNA and mRNA, respectively). After synthesis the number of templates is not changed. The corresponding reaction is therefore:

$$A \xrightarrow{k} A + B$$

Molecule A produces molecule B at a rate k (in units of $(\text{concentration} \times \text{time})^{-1}$). The master equation describes how the probability to be in state [n_1, n_2] (n_1 A molecules, n_2 B molecules) at time t changes in time. For the reaction above:

$$\dot{p}(n_1, n_2, t) = -kn_1 p(n_1, n_2, t) + kn_1 p(n_1, n_2 - 1, t)$$

The first term reflects a transition from state [n_1, n_2] to state [n_1, n_2+1] and therefore leads to a decrease in $p(n_1, n_2, t)$. The second term denotes the transition [n_1, n_2-1] $\rightarrow$ [n_1, n_2] and leads to an increased $p(n_1, n_2, t)$. The master equation above is linear and can be solved for the moments by constructing the moment generating function. In general for N system variables the moment generating function is given by:

$$F(z_1, z_2, ..., z_N, t) = \sum_{n_1, n_2, ..., n_N} z_1^{n_1} z_2^{n_2} ... z_N^{n_N} p(n_1, n_2, ..., n_N, t)$$

,

where the sum runs over all possible states for each n_i (in this case, from 0 to ∞). This function has the following useful properties:

$$F\big|_1 = 1, \quad \left.\frac{\partial F}{\partial z_i}\right|_1 = \langle n_i \rangle, \quad \left.\frac{\partial^2 F}{\partial z_i^2}\right|_1 = \langle n_i (n_i - 1) \rangle, \quad \left.\frac{\partial^2 F}{\partial z_i \partial z_j}\right|_1 = \langle n_i n_j \rangle,$$

where $\big|_1$ means that the function is evaluated at $z_j = 1$ for all j. These expressions justify the name "moment generating": we can obtain the moments of the probability distribution by evaluating the partial derivatives of the function.

Multiplying the master equation above by $z_1^{n_1} z_2^{n_2}$ on both sides gives:

$$\sum_{n_1,n_2} z_1^{n_1} z_2^{n_2} \dot{p}(n_1,n_2,t) = -k \sum_{n_1,n_2} z_1^{n_1} z_2^{n_2} n_1 p(n_1,n_2,t) + k \sum_{n_1,n_2} z_1^{n_1} z_2^{n_2} n_1 p(n_1,n_2-1,t)$$

This equation can be simplified significantly by realizing that:

$$\frac{\partial F}{\partial z_1} = \sum_{n_1} \left(n_1 z_1^{n_1 - 1}\right) p(n_1,n_2,t) \quad \Rightarrow \quad z_1 \frac{\partial F}{\partial z_1} = \sum_{n_1} n_1 z_1^{n_1} p(n_1,n_2,t)$$

$$\sum_{n_1=0,n_2=0} n_1 z_1^{n_1} z_2^{n_2} p(n_1,n_2-1,t) = z_1 \sum_{n_1=0,n_2=-1} \left(n_1 z_1^{n_1-1}\right) z_2^{n_2} p(n_1,n_2-1,t)$$

$$= z_1 \sum_{n_1=0,n'_2=0} \left(n_1 z_1^{n_1-1}\right) z_2^{n'_2+1} p(n_1,n'_2,t) = z_1 z_2 \frac{\partial F}{\partial z_1}$$

where the change in the lower limit of the sum for n_2 is allowed because $p(n_1,-1,t) = 0$. This leads to:

$$\dot{F}(z_1,z_2,t) = k z_1 (z_2 - 1) \frac{\partial F}{\partial z_1}$$

In the special case of synthesis from a fixed number of templates ($n_1 = n$), the equation for the moment generating function reduces to:

$$\dot{F}(z_2,t) = kn(z_2 - 1)F$$

This equation can be explicitly solved, but in itself it does not represent the full process. We therefore will obtain the expressions for the other terms before combining them to model a real situation.

Degradation

Now consider the degradation reaction:

$$B \xrightarrow{\gamma} 0 .$$

This reaction can represent two different processes: degradation, where molecule B is converted into a species which is not part of the subset of interest, and dilution, where it is physically separated from the volume of interest. In latter context γ is the degradation rate and $\ln(2)/\gamma$ the half-life of the molecule. The master equation for this reaction is:

$$\dot{p}(n_1,t) = -\gamma n_1 p(n_1,t) + \gamma(n_1+1)p(n_1+1,t)$$

Using the same strategy as above the time evolution of the moment generation function yields:

$$\dot{F}(z_1,t) = -\gamma(z_1-1)\frac{\partial F}{\partial z_1}$$

Forward reaction, conservation of total number of molecules

Now consider the reaction:

$$A \xrightarrow{k} B$$

where $n_o + n_1 = n = const.$ Since the total number n is conserved, the system is defined by only one variable. We will use n_2 as the single variable to define this system. For the reaction above:

$$\dot{p}(n_2,t) = -k(n-n_2)p(n_2,t) + k(n-n_2+1)p(n_2-1,t)$$

This leads[1] to:

$$\dot{F}(z_2,t) = -knF + kz_2\frac{\partial F}{\partial z_2} + k(n+1)z_2F - kz_2\left(z_2\frac{\partial F}{\partial z_2} + F\right) = kn(z_2-1)F - kz_2(z_2-1)\frac{\partial F}{\partial z_1}$$

Based on these elementary reactions larger chemical networks can be built up. The results above are summarized in Table 1.

[1] In this case, the sums only go up to n, instead of ∞. However, the extra terms that appear when applying the change of variables cancel with each other.

	Reaction Type	$\dot{F} =$
I	$A \xrightarrow{k} A + B$	$kz_1(z_2 - 1)\frac{\partial F}{\partial z_1}$
II	$B \xrightarrow{\gamma} 0$	$-\gamma(z_1 - 1)\frac{\partial F}{\partial z_1}$
III	$A \xrightarrow{k} B$ ($n_1 + n_2 = n$ = const., in terms of n_2)	$kn(z_2 - 1)F - kz_2(z_2 - 1)\frac{\partial F}{\partial z_2}$
.IV	$A \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} B$	$k_1(z_2 - z_1)\frac{\partial F}{\partial z_1} + k_{-1}(z_1 - z_2)\frac{\partial F}{\partial z_2}$

Table 1. *Moment generating function equations for elementary reactions. The master equation for each reaction type produces different terms which can be combined to model more complex processes.*

Noise properties of a constitutively expressed gene

Based on the results for these elementary reactions the equation for the moment generating functions of more complex networks can be easily deduced. First let us consider a constitutive expressed gene in a single copy in the chromosome of a bacterium. In this case the state of this system at any time is defined by the number of mRNA molecules *r* and number of proteins *p* for that gene. mRNA molecules are synthesized off the template DNA strand at a rate k_R and are translated at a rate k_P. The mRNA and protein degradation are described by the destruction rates γ_R and γ_P respectively (Fig. 1).

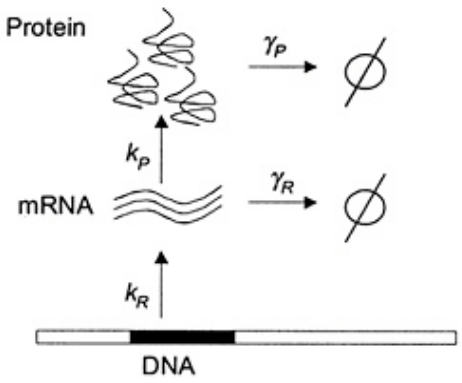

Figure 1*. Basic model for constitutive expression of a single gene. Only four individual reactions are considered: creation of mRNA from a DNA template, creation of proteins from individual mRNA molecules, and the degradation/dilution of both species.*

Based on the results in Table 1 the moment generating function can be deduced directly:

$$\dot{F}(z_1,z_2,t) = k_R(z_1-1)F + k_P z_1(z_2-1)\frac{\partial F}{\partial z_1} - \gamma_R(z_1-1)\frac{\partial F}{\partial z_1} - \gamma_P(z_2-1)\frac{\partial F}{\partial z_2}$$

The first two terms are the transcription and translation reactions (Table 1, type I) and the last two terms model degradation of mRNA and proteins respectively (Table 1, type II). Below the equation will be solved for the moments in the steady state ($\dot{F}=0$). In this case:

$$k_R(1-z_1)F = k_P z_1(z_2-1)\frac{\partial F}{\partial z_1} - \gamma_R(z_1-1)\frac{\partial F}{\partial z_1} - \gamma_P(z_2-1)\frac{\partial F}{\partial z_2}$$

The mean mRNA level $\langle r \rangle$ and protein level $\langle p \rangle$ are found by taking the derivative with respect to z_1 and z_2 respectively:

$$k_R(1-z_1)\frac{\partial F}{\partial z_1} - k_R F = \left[k_P z_1(z_2-1) - \gamma_R(z_1-1)\right]\frac{\partial^2 F}{\partial z_1^2} + \left[k_P(z_2-1) - \gamma_R\right]\frac{\partial F}{\partial z_1} - \gamma_P(z_2-1)\frac{\partial^2 F}{\partial z_2 \partial z_1}$$

evaluating both expressions at $z_1 = z_2 = 1$ gives:

$$\langle r \rangle = \frac{k_R}{\gamma_R}$$
$$\langle p \rangle = \frac{k_R k_P}{\gamma_R \gamma_P}$$

The results are consistent with the equivalent deterministic system:

$$\langle \dot{r} \rangle = k_R - \gamma_R \langle r \rangle$$
$$\langle \dot{p} \rangle = k_P \langle r \rangle - \gamma_P \langle p \rangle$$

The fluctuations in mRNA and proteins level are found by differentiating the above equations again with respect to z_1 and z_2 and evaluating at $z_1 = z_2 = 1$:

$$\langle r^2 \rangle - \langle r \rangle^2 = \langle r \rangle$$
$$\langle rp \rangle - \langle r \rangle \langle p \rangle = \frac{\langle p \rangle}{1+\gamma_R/\gamma_P}$$
$$\langle p^2 \rangle - \langle p \rangle^2 = \langle p \rangle \left(\frac{k_P/\gamma_R}{1+\gamma_P/\gamma_R} + 1 \right)$$

Further moments can be obtained sequentially in this manner. Note that for a random variable with Poissonian distribution, all moments are equal, so the variance over the mean equals one, as is the case for the mRNA in this model. The protein number fluctuates with a higher than poissonian noise, the correction determined primarily by the term k_P / γ_R (the "burst size"), which corresponds to the average number of proteins produced per mRNA [9].

In simple cases like this, the moments can also be obtained as a function of time. For a single gene, the noise out of equilibrium can be 40% larger than its steady state value in the limit of short mRNA lifetimes [9]. A more detailed modeling of this process could include more intermediate detailed processes, such as the random steps that a ribosome takes along an mRNA, but most turn out to have little effect when compared in simulations. However, when a repressor or activator is present, its binding and unbinding might have to be included in the model, for this can be a major source of noise. It is in this context that the terms shown in Table 1, type III, are needed. Furthermore, the repressor concentration itself might be fluctuating, in which case we have to consider the entire system of genes.

Linearized matrix formulation

The method above can also be used for interacting systems of genes, but solving it is not straightforward unless the connections are linear. Alternatively, if the system is at a stable point in steady state, the interaction can be linearized around the steady state value. A
practical way of writing this out is in matrix form. The transition probabilities for species x_i are given by $f_i(x_1,x_2,...,x_n)$ for creation and γ_i for destruction, and A and Γ are the matrices defined by $A_{ij} = \left.\frac{\partial f_i}{\partial x_j}\right|_{\langle x_1\rangle,\langle x_2\rangle...}$ and $\Gamma_{ij} = \gamma_i \delta_{ij}$. Letting x be the vector of chemical species, the linearized macroscopic equations are then given by $\langle \dot{x} \rangle = (A - \Gamma)\langle x \rangle$.
Note that since in many cases the macroscopic equations include constant creation terms. If the system is linear it might be necessary to include an additional variable, which is not fluctuating and allows the inclusion of the constant terms in the compact matrix form. As an illustration of this, the matrices for the single gene case are

$$A = \begin{bmatrix} 0 & 0 & 0 \\ k_R & 0 & 0 \\ 0 & k_P & 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \gamma_R & 0 \\ 0 & 0 & \gamma_P \end{bmatrix}.$$

where the state vector is $x^T = (d, r, p)$ where d is the gene copy number. This constant state coordinate needs not to represent an actual chemical; for a system where many species have a constant creation rate, these rates can all be placed in the first column of A (setting d=1 and $\Gamma_{1j} = \Gamma_{j1} = 0$). An example of this is the matrix for the case of two

interacting genes, linearized around steady state, with fixed gene copy numbers d_1 and d_2 respectively, and where the first gene (r_1,p_1) represses the second (r_2,p_2) with transfer function $f(p_1)$:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ d_1 k_{R1} & 0 & 0 & 0 & 0 \\ 0 & k_{P1} & 0 & 0 & 0 \\ d_2 k_{R2} & 0 & d_2 \left.\frac{\partial f}{\partial p_1}\right|_{\langle p_1 \rangle} & 0 & 0 \\ 0 & 0 & 0 & k_{P2} & 0 \end{bmatrix}, \quad \Gamma = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \gamma_{R1} & 0 & 0 & 0 \\ 0 & 0 & \gamma_{P1} & 0 & 0 \\ 0 & 0 & 0 & \gamma_{R2} & 0 \\ 0 & 0 & 0 & 0 & \gamma_{P2} \end{bmatrix},$$

where $k_{R2} = f\left(\langle p_1 \rangle\right) - \left.\frac{\partial f}{\partial p_1}\right|_{\langle p_1 \rangle} \langle p_1 \rangle$.

Written in terms of these matrices, the master equation in generating function form would be

$$\dot{F} = \sum_i (1 - z_i) \left(\Gamma_{ii} \frac{\partial F}{\partial z_i} - \sum_j A_{ij} z_j \frac{\partial F}{\partial z_j} \right).$$

At steady state, $\dot{F} = 0$, and taking the derivative with respect to z_l we obtain:

$$0 = \sum_i (1 - z_i) \left(\Gamma_{ii} \frac{\partial^2 F}{\partial z_i \partial z_l} - \sum_j A_{ij} z_j \frac{\partial^2 F}{\partial z_j \partial z_l} - A_{il} \frac{\partial F}{\partial z_l} \right) - \left(\Gamma_{ll} \frac{\partial F}{\partial z_l} - \sum_j A_{lj} z_j \frac{\partial F}{\partial z_j} \right)$$

Setting all z_i=1, we have for each i

$$0 = \Gamma_{ii} \left.\frac{\partial F}{\partial z_i}\right|_1 - \sum_j A_{ij} z_j \left.\frac{\partial F}{\partial z_j}\right|_1 \Rightarrow 0 = (\Gamma - A) \nabla F|_1 = (\Gamma - A) \langle x_1 \rangle,$$

corresponding to the macroscopic result. Similarly, differentiating again and evaluating at z_i=1,

$$\begin{aligned} 0 &= \left(\Gamma_{ii} \left.\frac{\partial^2 F}{\partial z_i \partial z_l}\right|_1 - \sum_j A_{ij} z_j \left.\frac{\partial^2 F}{\partial z_j \partial z_l}\right|_1 - A_{il} \left.\frac{\partial F}{\partial z_l}\right|_1 \right) + \left(\Gamma_{ll} \left.\frac{\partial^2 F}{\partial z_l \partial z_i}\right|_1 - \sum_j A_{lj} z_j \left.\frac{\partial^2 F}{\partial z_j \partial z_i}\right|_1 - A_{li} \left.\frac{\partial F}{\partial z_i}\right|_1 \right), \\ &= \left((\Gamma - A) \nabla \nabla^T F|_1 - A \Theta F|_1 \right) + \left((\Gamma - A) \nabla \nabla^T F|_1 - A \Theta F|_1 \right)^T \end{aligned}$$

where $\Theta_{ij} = \delta_{ij} \frac{\partial}{\partial z_i}$. These linear equations can be solved for the means, variances and correlations.

This approach is very general and the resulting matrix equations can be solved directly[2]. However, even for the case of just two interacting genes this requires a 5x5 matrix system as shown, so it gets cumbersome for larger systems even though most entries are zero. Using symbolic matrix manipulation software it is straightforward to obtain the desired expressions, so for known parameters this is a good method for obtaining values without further approximations.

2. THE LANGEVIN APPROACH

An alternate approach that allows for a more straightforward interpretation and scales easily to different levels of detail is the use of a Langevin equation.
The Langevin approach consists essentially of adding a noise term to the deterministic equations. This noise term can represent the effect of the intrinsic fluctuations [20] or the external inputs of the system [21].
For *x*, the concentration of some chemical species,

$$\dot{x} = f(x) \rightarrow \dot{x} = f(x) + q(x)\varepsilon(t)\,,$$

where the random variable $\varepsilon(t)$ is determined by its statistical properties. Formally, this can be any random process, but in practice we assume white-noise statistics, which will give approximate values for the first two moments. The conditions for white noise are:

$$\langle \varepsilon(t) \rangle = 0, \qquad \langle \varepsilon(t)\varepsilon(t+\tau) \rangle = \delta(\tau),$$

where $\langle\ \rangle$ denotes an ensemble average. Since we are interested in the steady state fluctuations, we will assume the coefficient of the noise term to be constant[3], i.e. evaluated at $\langle x \rangle_{ss}$.
For the case of our basic model of the single gene, we have two macroscopic equations representing mRNA and protein creation, respectively:

$$\dot{r} = k_R - \gamma_R r \rightarrow \dot{r} = k_R - \gamma_R r + q_r \varepsilon_r$$
$$\dot{p} = k_P r - \gamma_P p \rightarrow \dot{p} = k_P r - \gamma_P p + q_p \varepsilon_p\,,$$

[2] This can be summarized in a very practical way [13] in terms of the logarithmic gains to obtain an equation which reflects the resulting components of the noise.
[3] For the case where *q(x)* is not constant, the stochastic differential equation will be understood to follow the Stratonovich interpretation [19,22]. This allows a general Fokker-Planck equation to be written in this form, but will not be necessary in the cases of interest.

where the coefficients of the noise terms are to be determined. Clearly, $\langle r \rangle = k_R / \gamma_R$ and $\langle p \rangle = k_P \langle r \rangle / \gamma_P$ from the condition of zero mean for the noise term. The difference with the steady state $\delta r = r - \langle r \rangle$ follows the equation

$$\delta\dot{r} + \gamma_R \delta r = q_r \varepsilon_r$$

Fourier transforming, we obtain

$$\delta\hat{r}(\omega) = \frac{q_r \hat{\varepsilon}_r}{i\omega + \gamma_R},$$

so after multiplying by the complex conjugate and taking the average,

$$\left\langle \left| \delta\hat{r}(\omega) \right|^2 \right\rangle = \frac{{q_r}^2}{\omega^2 + \gamma_R^2}$$

The steady state fluctuations are given by the inverse Fourier transform with $t = 0$[4]:

$$\delta r^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{{q_r}^2}{\omega^2 + \gamma_R^2} d\omega = \frac{{q_r}^2}{2\pi\gamma_R} \int_{-\infty}^{\infty} \frac{dx}{x^2 + 1} = \frac{{q_r}^2}{2\gamma_R}.$$

But since the production of mRNA is in this model a single step, independent random process, it has a Poisson distribution, so the variance equals the mean, which implies

$$\frac{{q_r}^2}{2\gamma_R} = \frac{k_R}{\gamma_R} \Rightarrow {q_r}^2 = 2k_R.$$

For the number of proteins, we have

$$\delta\dot{p} + \gamma_P \delta p = \delta r + q_p \varepsilon_p,$$
$$\delta\hat{p}(\omega) = \frac{\delta\hat{r}(w) + q_p \hat{\varepsilon}_p}{i\omega + \gamma_P},$$

but in this case we also need to notice that $\left\langle \delta\hat{r}(w) \hat{\varepsilon}_p^* \right\rangle = \left\langle \delta\hat{r}(w) \right\rangle \left\langle \hat{\varepsilon}_p^* \right\rangle = 0$, since these are two independent random processes with zero mean. So in this case,

$$\left\langle \left| \delta\hat{p}(\omega) \right|^2 \right\rangle = \frac{\left\langle \left| \delta\hat{r}(\omega) \right|^2 \right\rangle + {q_p}^2}{\omega^2 + \gamma_P^2} = \frac{{q_r}^2}{\left(\omega^2 + \gamma_R^2\right)\left(\omega^2 + \gamma_P^2\right)} + \frac{{q_p}^2}{\omega^2 + \gamma_P^2}.$$

[4] From the Wiener-Khintchine theorem; see [23].

Using $q_r^{\,2} = 2k_R$ and $q_p^{\,2} = 2k_p \dfrac{k_R}{\gamma_R}$, (since this represents the internal noise and for a fixed number of mRNAs the production of proteins is also a Poissonian process). Performing the inverse transform[5],

$$\delta p^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{q_r^{\,2}}{\left(\omega^2+\gamma_R^2\right)\left(\omega^2+\gamma_P^2\right)} + \frac{q_p^{\,2}}{\omega^2+\gamma_P^2}\, d\omega = \frac{k_P^{\,2}k_R}{\left(\gamma_R^{\,2}-\gamma_P^{\,2}\right)}\left[\frac{1}{\gamma_P}-\frac{1}{\gamma_R}\right] + \frac{k_P k_R}{\gamma_P\gamma_R}$$

for comparison with the previous result, note that $\langle p\rangle = \dfrac{k_P k_R}{\gamma_P\gamma_R}$, so this can be rewritten as

$$\delta p^2 = \langle p\rangle\left(\frac{k_P/\gamma_R}{1+\gamma_P/\gamma_R}+1\right)$$

This is identical to the result obtained by the master equation. This method can be readily generalized for many interacting genes when the system is fluctuating around a steady state. As an example, we will analyze the case where one gene represses a second gene. Let y_0, y_1 be the protein numbers of each gene, and let $f(y_0)$ be the rate of creation of the second protein as a function of the first. This means that the equations describing this system are

$$\dot{y}_0 = k_0 - \gamma_0 y_0\,,$$
$$\dot{y}_1 = f(y_0) - \gamma_1 y_1\,.$$

Note that the equations include the entire process of producing a protein, so mRNA levels are no longer explicitly calculated. Including the Langevin noise term and looking at the fluctuations from steady state,

$$\delta\dot{y}_0 = -\gamma_0\delta y_0 + q_0\varepsilon_0\,,$$
$$\delta\dot{y}_1 = f(y_0) - f(\langle y_0\rangle) - \gamma_1\delta y_1 + q_1\varepsilon_1 \approx c_0\delta y_0 - \gamma_1\delta y_1 + q_1\varepsilon_1\,,$$

where $c_0 = \left.\dfrac{df}{dy_0}\right|_{\langle y_0\rangle}$ and each noise term has the same conditions as before. This linearization is valid at each stable point, but not for transitions between different stable points or for limit cycles. For very small numbers n of chemicals this also breaks down, because since this processes are mostly Poissonian, the fluctuations are of order $\sqrt{n}$ so a Taylor expansion might not be valid. Fourier transforming and taking the square and the average as before, we get

---

[5] $\dfrac{1}{2\pi}\int \dfrac{d\omega}{\left(\omega^2+\gamma^2\right)^n} = \dfrac{1}{2\sqrt{\pi}\gamma^{2n-1}}\dfrac{\Gamma(n-1/2)}{\Gamma(n)}$, where $\Gamma(n)=\Gamma(n-1)(n-1)$, $\Gamma(1)=1$ and $\Gamma(1/2)=\sqrt{\pi}$.

$$\left\langle \left|\delta\hat{y}_0(\omega)\right|^2 \right\rangle = \frac{{q_0}^2}{\omega^2+\gamma_0^2}$$

$$\left\langle \left|\delta\hat{y}_1(\omega)\right|^2 \right\rangle = \frac{c_0^2\left\langle \left|\delta\hat{y}_0(\omega)\right|^2 \right\rangle + {q_1}^2}{\omega^2+\gamma_1^2} = \frac{c_0^2{q_0}^2}{\left(\omega^2+\gamma_1^2\right)\left(\omega^2+\gamma_0^2\right)} + \frac{{q_1}^2}{\omega^2+\gamma_1^2}$$

The correlation between the genes can also be calculated, from

$$\left\langle \delta\hat{y}_1\delta\hat{y}_0^* \right\rangle = \left\langle \left(\frac{c_0\delta\hat{y}_0(w)+q_1\hat{\varepsilon}_1}{i\omega+\gamma_1} \right)\delta\hat{y}_0^*(w) \right\rangle = \frac{c_0\left\langle \left|\delta\hat{y}_0(\omega)\right|^2 \right\rangle}{i\omega+\gamma_1} = \frac{c_0{q_0}^2}{\left(i\omega+\gamma_1\right)\left(\omega^2+\gamma_0^2\right)},$$

where the term $\left\langle \hat{\varepsilon}_1\delta\hat{y}_0^* \right\rangle$ vanishes because the fluctuations in the first gene are independent from the internal fluctuations in the second gene. In many cases, the decay time will be determined primarily by the dilution time, so it will be the same for all genes. This assumption simplifies the expressions that are obtained upon transforming back:

$$\left\langle \delta y_0^2 \right\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{{q_0}^2}{\omega^2+\gamma^2}d\omega = \frac{{q_0}^2}{2\gamma}$$

$$\left\langle \delta y_1^2 \right\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{c_0^2{q_0}^2}{\left(\omega^2+\gamma^2\right)^2} + \frac{{q_1}^2}{\omega^2+\gamma^2}d\omega = \frac{c_0^2{q_0}^2}{4\gamma^3} + \frac{{q_1}^2}{2\gamma}$$

$$\left\langle \delta y_1\delta y_0 \right\rangle = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{c_0{q_0}^2}{\left(i\omega+\gamma\right)\left(\omega^2+\gamma^2\right)}d\omega = \frac{c_0{q_0}^2}{2\pi}\int_{-\infty}^{\infty}\frac{\left(i\omega-\gamma\right)}{\left(\omega^2+\gamma^2\right)^2}d\omega = \frac{c_0{q_0}^2}{4\gamma^2},$$

where the irrational part of the integral vanishes because of parity. From our previous results we know that for a single gene,

$$\left\langle \delta y_0^2 \right\rangle = \frac{{q_0}^2}{2\gamma} = \left\langle y_0 \right\rangle\left(\frac{k_{Po}/\gamma_{Ro}}{1+\gamma/\gamma_{Ro}} + 1 \right) \approx \left\langle y_0 \right\rangle\left(b_0+1\right) \Rightarrow {q_0}^2 \approx 2\gamma\left\langle y_0 \right\rangle\left(b_0+1\right),$$

where b_i is the burst size for gene *i.* For basic Hill-type repression,

$$f\left(y_0\right) = \frac{k_1}{1+\left(y_0/Y_{1/2}\right)^h} + B_1 \Rightarrow c_0 = \frac{-k_1}{\left(1+\left(y_0/Y_{1/2}\right)^h\right)^2}\left(\frac{\left\langle y_0 \right\rangle}{Y_{1/2}}\right)^{h-1}\frac{h}{Y_{1/2}} = \left\langle y_1 \right\rangle^2\left(\frac{\left\langle y_0 \right\rangle}{Y_{1/2}}\right)^{h-1}\frac{-h\gamma^2}{k_1Y_{1/2}}$$

where k_1+B_1 is the maximum creation rate, $Y_{1/2}$ is the half induction point, h is the Hill coefficient and B_1 is the basal transcription level. Assuming that the internal noise for the second gene alone has the same form,

$$q_1^{\,2} \approx 2\gamma \langle y_1 \rangle (b_1 + 1),$$

the variance and correlation can be explicitly written as

$$\left\langle \delta y_1^2 \right\rangle = \left(\frac{\langle y_0 \rangle}{Y_{1/2}} \right)^{2h} \left(\frac{h\gamma}{2k_1} \right)^2 \frac{\langle y_1 \rangle^2}{\langle y_0 \rangle} 2(b_0 + 1) + \langle y_1 \rangle (b_1 + 1)$$

$$\left\langle \delta y_1 \delta y_0 \right\rangle = \langle y_1 \rangle^2 \left(\frac{\langle y_0 \rangle}{Y_{1/2}} \right)^h \frac{-h\gamma}{2k_1} (b_0 + 1).$$

Note that we need the parameters of the macroscopic equations plus an "internal" parameter for each gene, $b_i = k_{Pi} / \gamma_{Ri}$ which depends on the parameters of the macroscopic equations for each gene.