About this document
In this document, we provide detailed derivations and analyses pertaining to the results presented in the main text.

We aim to describe the stochastic behavior of a cell under conditions of steady-state exponential growth. A cell contains many protein species. Each of these may affect growth to some extent via an effect on metabolism. (Note that, here and below, we use the term metabolism in a broad sense; it is intended to encompass all catabolic and anabolic processes required for biomass production and cell growth, including protein synthesis.) As mentioned in the main text, our results are based on several crucial assumptions. In particular, we assume that

- the protein density of the cell is under tight homeostatic control,
- protein synthesis is inherently stochastic,
- the cellular growth rate is an intensive quantity.

These assumptions are used repeatedly in the text below.

Derivation of the linear noise model
We start with a detailed derivation of the linear noise model.

Abundances, volume, and growth rate
The abundances of the protein species are given by the vector \( n = \{n_1(t), n_2(t), \ldots \} \), where \( n_i(t) \) is the copy number of protein species \( i \) at time \( t \). We define the total protein mass of the cell by

\[
M = \sum_i n_i ,
\]

where we ignore that different proteins have different molecular weights. (To avoid clutter, we left out the explicit time dependence of the variables; below, we do so whenever this is unlikely to cause confusion.) Under the assumption that the protein density of a cell is tightly controlled, \( M \) is proportional to the volume of the cell. Therefore, we define the instantaneous growth rate \( \mu \) as the rate at which \( M \) increases:

\[
\mu \equiv \frac{1}{M} \frac{\mathrm{d}M}{\mathrm{d}t} .
\]
Proteome fractions and their dynamics

The proteome fractions of the different protein species in the cell are given by the vector

\[ \phi := \frac{n}{M} \]  

(3)

Again, assuming that the mass \( M \) is proportional to the cellular volume, we can interpret the mass fractions as concentrations.

We denote the rate of protein synthesis by the vector \( p(t) = (p_1(t), p_2(t), \ldots) \) and assume that, under conditions of balanced exponential growth, degradation of proteins is negligible compared to protein dilution due to growth [40]. Then the time derivative of equation (3) yields the following system of dynamical equations:

\[ \frac{d\phi}{dt} = \pi - \mu \phi, \]  

(4)

where the vector \( \pi := p/M \) is the protein synthesis rate per unit of protein mass.

By definition, the mass fractions add up to one:

\[ \sum_i \phi_i = 1. \]  

(5)

Therefore, an increase in the concentration of one protein must go at the expense of the concentration of another. Applied to equation (4), this constraint also implies that

\[ \mu = \sum_i \pi_i. \]  

(6)

That is, the growth rate is equal to the total rate of protein synthesis per unit of mass.

In reality, even if protein density is under strict control, it is likely to show fluctuations on short time scales. In mammalian cells, mass fluctuations have indeed been measured, but appeared to dissipate on time scales of less than 20 s [55]. By comparison, the auto-correlation times of the expression and growth-rate fluctuations observed in E. coli cells are very long—they are of the order of hours. On this time scale, we therefore expect equation (6) to be a reasonable approximation.

Stochastic protein-synthesis rate

The protein synthesis rate of each protein species \( i \) is assumed to have the following form:

\[ \pi_i = f_i J(n) / M + N_i \]  

(7)

The first term on the right-hand side models the deterministic dependence of the protein synthesis rate on the protein content of the cell. The function \( J(n) \) represents the global metabolic flux that fuels protein synthesis and growth. The vector \( f = \{f_1, f_2, \ldots\} \) indicates which fraction of this metabolic flux is allocated towards the synthesis of each of the protein species.

As such, it obeys the constraint

\[ \sum_i f_i = 1. \]  

(8)

In general, \( f \) is itself likely to be a function of intra- and extracellular variables, including the protein content of the cell, due to the dynamical regulation of gene expression. Further below we will reflect on this general case, but, unless indicated otherwise, we focus on the special case where \( f \) is constant.

The second term on the right-hand size of equation (7) models the stochasticity of protein synthesis. Each \( N_i \) is a colored Ornstein–Uhlenbeck noise source with zero mean, reversion rate \( \beta_i \), and amplitude \( \theta_i \), obeying the following stochastic differential equation (Itô):

\[ dN_i(t) = -\beta_i N_i dt + \theta_i dW_i(t), \]  

(9)

where each \( W_i(t) \) is a Wiener process. Importantly, all noise sources are assumed to be mutually independent.

Stochastic growth

Combined, equations (6) and (7) yield the following expression for the growth rate \( \mu \):

\[ \mu = J(n)/M + \sum_i N_i. \]  

(10)

The first term, \( J(n)/M \), is a function of the protein content of the cell; we call it \( \mu_d(n) \). The assumption that the growth rate is an intensive quantity translates to the requirement that

\[ \mu_d(\alpha n) = \mu_d(n), \]  

(11)

for any \( \alpha > 0 \). Because \( \mu_d \) is intensive and \( M \) is extensive, the metabolic flux \( J(n) = M \mu_d(n) \) must be extensive, \( i.e. \):

\[ J(\alpha n) = \alpha J(n), \]  

(12)

for any \( \alpha > 0 \). This is indeed routinely assumed in the field of metabolic control analysis [64]. Using equation (11), we note that \( \mu_d(n) = \mu_d(n/M) = \mu_d(\phi) \), so that equations (7) and (10) can be rewritten as

\[ \pi_i = f_i \mu_d(\phi) + N_i, \]  

(13)

\[ \mu = \mu_d(\phi) + \sum_i N_i. \]  

(14)

Equation (14) shows that all noise sources affect the growth rate in two ways: directly through the second term, and indirectly through the first, because of fluctuations in the composition \( \phi \).

Provided a function \( \mu_d(\phi) \) is given, equations (4), (13) and (14) together fully specify the dynamics of \( \pi, \phi, \) and \( \mu \).
**Linearization**

To obtain analytical results, we assumed that fluctuations in the concentrations $\phi$ are small enough to warrant a linear approximation. Below, we describe the linearization of equations (4), (13), and (14) around the time average $\phi_0$ of $\phi$.

We start with equation (4). We introduce transformed variables:

$$\delta \phi := \phi - \phi_0,$$

$$\delta \pi := \pi - \pi_0,$$

$$\delta \mu := \mu - \mu_0,$$  

where $\mu_0 := \mu_0(\phi_0)$ and $\pi_0 := f \mu_0$. In terms of these variables, equation (4) reads:

$$\frac{\delta \phi_i}{\mu_0 \phi_{0,i}} = \left( \frac{f_i}{\phi_{0,i}} - 1 \right) + \left( \frac{f_i}{\phi_{0,i}} \right) \frac{\delta \pi_i}{\pi_0},$$

$$- \frac{\delta \mu_i}{\mu_0} - \frac{\delta \phi_i}{\phi_{0,i}} - \frac{\delta \phi_i}{\phi_{0,i}} \frac{\delta \mu_i}{\mu_0}. \tag{18}$$

To proceed, we exploit that $f_i = \phi_{0,i}$ to linear order. To see why, we first observe from equation (13) that

$$f_i = \frac{\mathbb{E} [\pi_i]}{\mathbb{E} [\mu_0(\phi)]}, \tag{19}$$

where $\mathbb{E} [X]$ denotes the steady-state expectation value of variable $X$. Next, equation (14) implies that

$$\mathbb{E} [\mu_0(\phi)] = \mathbb{E} [\mu], \tag{20}$$

and equation (4) that

$$\mathbb{E} [\pi_i] = \mathbb{E} [\mu \phi_i]. \tag{21}$$

Therefore,

$$f_i = \frac{\mathbb{E} [\mu \phi_i]}{\phi_{0,i} \mathbb{E} [\mu]} = 1 + \text{Cov} \left( \delta \phi_i, \frac{\delta \mu_i}{\mathbb{E} [\mu]} \right), \tag{22}$$

which is 1 up to quadratic corrections. We can estimate the size of these corrections (the covariance term) as follows. We write:

$$\text{Cov} \left( \delta \phi_i, \frac{\delta \mu_i}{\mathbb{E} [\mu]} \right) = R \eta_\phi \eta_\mu, \tag{23}$$

where $R$ is the correlation coefficient between $\phi_i$ and $\mu_i$, and $n_i$ is the coefficient of variation of variable $x$. Taking rather large estimates for each, we arrive at $R \eta_\phi \eta_\mu \approx 0.5 \times 0.4 \times 0.25 = 0.05$. This shows that quadratic corrections are small and we can indeed assume that $f_i = \phi_{0,i}$ to good approximation. (We also verified this using simulations; see below.) Based on this, equation (18) takes the following form after linearization:

$$\frac{\delta \phi_i}{\mu_0 \phi_{0,i}} = \frac{\delta \pi_i}{\pi_0} - \frac{\delta \mu_i}{\phi_{0,i}} - \frac{\delta \phi_i}{\phi_{0,i}} \frac{\delta \mu_i}{\mu_0}. \tag{24}$$

We now move to equation (13). After linearization, it can be written as:

$$\frac{\delta \pi_i}{\pi_{0,i}} = \sum_j C_{ij} \frac{\delta \phi_j}{\phi_{0,j}} + v_i, \tag{25}$$

where we introduce growth-control coefficients (GCCs), defined as

$$C_{ij} := \left[ \frac{\phi_i}{\mu_0} \frac{\partial \mu_i}{\partial \phi_j} \right]_{\phi_0}. \tag{26}$$

(See p. 7 of this text for a further discussion of these quantities.) The Ornstein–Uhlenbeck noise terms $v_i := N_i/\pi_{0,i}$ are identical to the original noise terms $N_i$, except that their noise amplitudes $\theta_i$ are rescaled as $\theta_i = \theta_i/\pi_{0,i}$. Equation (25) shows that the synthesis of a protein species is affected by its own noise source, but also by fluctuations in the concentrations of all other proteins, in as far as these fluctuations affect the growth.

These same methods can be applied to equation (14), the linearized version of which is given by

$$\frac{\delta \mu_i}{\mu_0} = \sum_j C_{ij} \frac{\delta \phi_j}{\phi_{0,j}} + \mathcal{N}. \tag{27}$$

Here we introduced the shorthand

$$\mathcal{N} = \sum_i f_i v_i, \tag{28}$$

which is a weighted sum of all noise sources.

We now use equations (25) and (27) to rewrite equation (24) as

$$\frac{\delta \phi_i}{\mu_0 \phi_{0,i}} = v_i - \mathcal{N} - \frac{\delta \phi_i}{\phi_{0,i}}. \tag{29}$$

Equations (25), (27), and (29) together form the basis for the analyses in the main text.

To test the quality of the analytical results based on the linearization of the equations and the approximation that $f_i = \phi_{0,i}$, we simulated the nonlinear equations (4), (13) and (14) for a system containing just two proteins X and Y, using a convenient nonlinear function $\mu_0(\phi) = (a_X \phi_X + a_Y \phi_Y)/(\phi_X + \phi_Y)$. (Note that $\mu_0$ is indeed an intensive function, as it should be. See p. 9 of this document for details on the simulation method.) We then numerically calculated the cross-correlations based on these simulations, and compared them to the corresponding analytical predictions. We chose the parameters of the simulation (including $a_X$ and $a_Y$) such as to match the parameters used in Fig. 3AB of the main text. In that figure, the simulation results based on the non-linear equations are plotted (gray diamonds) alongside the analytical predictions based on the linearized equations (solid gray line). They show excellent agreement, despite significant noise levels (coefficients of variation of $\phi_X$, $\phi_Y$, and $\mu$ were 0.09, 0.19, and 0.26, respectively). This shows that the approximations made in this section are excellent for sufficiently smooth functions $\mu_0(\phi)$. 

---

3
Including dynamical gene regulation

In the previous sections, we assumed that a fixed fraction \( f_i \) of the metabolic flux is allocated towards the synthesis of protein \( i \). In this section we briefly describe the more general case where \( f \) depends on the composition of the cell, as expected due to transcriptional and translational regulation. For this case, analytical solutions are generally much harder to obtain. Nevertheless, we can gain some insight into the regulated system by studying the structure of the resulting equations.

We assume that \( f \) is an intensive quantity—that is, that it can be written as a function of the relative abundances \( \phi \). Then equation (13), for the production rate \( \pi_i \), generalizes to:

\[
\pi_i = f_i(\phi)\mu_0 + N_i,
\]

with \( 0 \leq f_i(\phi) \leq 1 \) and \( \sum_i f_i(\phi) = 1 \). This expression describes how the composition \( \phi \) of the cell differentially affects the production rate of each protein species. Linearizing equation (30) around \( \phi_0 \), we find:

\[
\delta \pi_i = \sum_i \frac{\partial}{\partial \phi_j} \left[ f_i(\phi)\mu_0 \phi_j \right] \delta \phi_j + N_i,
\]

\[
=f_i(\phi_0)\mu_0 \sum_j \left( \frac{\phi_{0,j}}{\mu_0} \frac{\partial \mu_0}{\partial \phi_j} + \frac{\phi_{0,j}}{f_i(\phi_0)} \frac{\partial f_i}{\partial \phi_j} \right) \delta \phi_j + N_i,
\]

which results in:

\[
\frac{\delta \pi_i}{\pi_{0,i}} = \sum_j \left( C_i^\mu + C_i^\phi \right) \frac{\delta \phi_j}{\phi_{0,j}} + \nu_i.
\]

Here, new control coefficients emerge: regulatory control coefficients (RCCs), defined as:

\[
C_i^\mu := \left[ \frac{\phi_j}{f_i(\phi_0)} \frac{\partial f_i}{\partial \phi_j} \right] \phi_0.
\]

These RCCs quantify to what extent small changes in the cell’s composition affect the allocation of the metabolic flux over the synthesis of various proteins species. As apparent from equation (32), the RCCs modulate the GCCs in an additive fashion.

Remarkably, the RCCs vanish from the expression for \( \delta \mu \):

\[
\delta \mu = \sum_i \pi_{0,i} \sum_j \left( C_i^{\mu} + C_i^{\phi} \right) \frac{\delta \phi_j}{\phi_{0,j}} + \sum_i \pi_{0,i} \nu_i
\]

\[
= \mu_0 \sum_j \left( C_i^{\mu} \right) \frac{\delta \phi_j}{\phi_{0,j}} + \mu_0 \sum_i \left( f_i(\phi_0) C_i^{\phi} \frac{\delta \phi_j}{\phi_{0,j}} + \mu_0 \phi_j \right)
\]

\[
= \mu_0 \sum_j \left( C_i^{\mu} \right) \frac{\delta \phi_j}{\phi_{0,j}} + \mu_0 \sum_j \left( \sum_i \frac{\delta f_i}{\delta \phi_j} \right) \delta \phi_j + \mu_0 \phi_j
\]

\[
= \mu_0 \left( \sum_j C_i^\mu \frac{\delta \phi_j}{\phi_{0,j}} + \phi_j \right).
\]

This derivation uses \( \pi_{0,i} = \mu_0 f_i(\phi_0) \) and hinges on the fact that \( \sum_i f_i(\phi) = 1 \), so that \( \sum_i \frac{\delta f_i}{\delta \phi_j} = 0 \).

The linearized dynamical equation for \( \delta \phi_i \) in the presence of regulation becomes:

\[
\frac{\delta \phi_i}{\phi_{0,i}} = \frac{\delta \pi_i}{\pi_{0,i}} - \frac{\delta \mu}{\mu_0} - \frac{\delta \phi_i}{\phi_{0,i}}
\]

\[
= \nu_i - N_i - \frac{\delta \phi_i}{\phi_{0,i}} + \sum_j C_j^\phi \frac{\delta \phi_j}{\phi_{0,j}}.
\]

By comparing with equation (29) we concluded that gene regulation adds a single term that describes (to linear order) the combined effect on \( f_i \) of the fluctuations in the concentrations of all proteins species \( j \).

Calculating statistical properties

Fourier transforms

The linearized expressions from p. 3 enable us to calculate various statistical properties. To do so, it is convenient to work with Fourier-transformed equations and variables. Based on equations (25), (27), and (29), we will now derive expressions for the Fourier transforms of the protein concentrations, the synthesis rates, and the growth rate. Further below, these expressions will be used to compute (co-)variances and correlations between the variables.

We denote the Fourier transform of variable \( y \) as \( \tilde{y} \). Taking the Fourier transform of both sides of equation (29) and using the transforms of equations (25) and (27), we find

\[
\tilde{\delta \phi_i} = \frac{\mu_0}{\mu_0 + i \omega} (\tilde{\nu}_j - \tilde{N}_j),
\]

where \( \omega \) is the Fourier frequency and \( i \) denotes the imaginary unit (to distinguish it from indices \( i \)). To arrive at this result, we used that

\[
\frac{\partial \tilde{f}_i}{\partial t} = i \omega \tilde{f}_i,
\]

for any time-dependent function \( f \).

By combining equation (36) with the Fourier transform of equation (27), we find a direct expression for the Fourier-transformed growth-rate deviations:

\[
\frac{\delta \mu}{\mu_0} = \sum_j \left( C_j^{\mu} \frac{\mu_0}{\mu_0 + i \omega} (\tilde{\nu}_j - \tilde{N}_j) + \tilde{N}_j \right)
\]

\[
= \sum_j \left( \frac{\mu_0}{\mu_0 + i \omega} C_j^{\mu} + \phi_{0,j} \right) \tilde{\nu}_j
\]

Here we used the definition of \( N \) (equation (28)), the fact that \( f_i = \phi_{0,i} \) to linear order, and the sum rule for GCCs presented in the main text:

\[
\sum_j C_i^\mu = 0.
\]
and the coefficients of variation (CV) are defined as
\[ \eta_x := \frac{\sqrt{\text{Var}(x)}}{\mathbb{E}[x]} = \frac{\sigma_x}{x_0}, \]
with \( \sigma_x \) the standard deviation of \( x \). The cross-correlation function \( R_{xy}(\tau) \) is given by
\[ R_{xy}(\tau) = \frac{\chi_{xy}(\tau)}{\sqrt{\text{Var}(x)\text{Var}(y)}}; \]
at delay \( \tau = 0 \) it reduces to the Pearson correlation coefficient.

The cross-covariances \( \mu, \phi_i \) and \( \pi_i \) are calculated by inserting (36), (38), and (40) into equation (42). In the process, we encounter products of noise sources \( \bar{v}_i^* \bar{v}_j \) (Fourier-transformed cross-covariances of the noise sources), which, based on equation (9), can be expressed as:
\[ \bar{v}_i^* \bar{v}_j = \delta_{ij} \theta_i^2 \frac{2}{\beta_i^2 + \omega^2}. \] Here, the Kronecker delta \( \delta_{ij} \) reflects the mutual independence of the noise sources. After using this identity, all cross-covariances can be written as linear combinations of three families of functions, \( B_i(\tau) \), \( A_i(\tau) \) and \( S_i(\tau) \), resulting from three Fourier integrals that can be solved explicitly with Cauchy’s residue theorem.

The first function is
\[ B_i(\tau) := \int e^{i\omega \tau} |\bar{v}_i|^2 d\omega = \frac{\theta_i^2}{2\beta_i} e^{-\beta_i |\tau|}. \] This function is symmetric; it is the auto-covariance of an Ornstein–Uhlenbeck process and therefore characterized by a single timescale \( \beta_i \).

The second function is
\[ A_i(\tau) := \int e^{i\omega \tau} \frac{\mu_0}{2\pi \mu_0 - i\omega} |\bar{v}_i|^2 d\omega \]
\[ = \theta_i^2 \mu_0 \begin{cases} e^{-\mu_0 |\tau|} & \text{if } \tau \leq 0, \\ e^{-\beta_i |\tau|} & \text{if } \tau \geq 0. \end{cases} \] This is an asymmetric function. It appears in the cross-correlation between variables \( X \) and \( Y \) when \( X \) is instantaneously affected by an Ornstein–Uhlenbeck noise source, while \( Y \) is affected by the same noise source with a delay characterize by rate \( \mu_0 \). For example, the effect of noise in the synthesis of a protein on its concentration involves such a delay; as a result, the function \( A_i(\tau) \) appears in the covariance between \( \pi_i \) and \( \phi_i \).

The third function is
\[ S_i(\tau) := \int e^{i\omega \tau} \frac{\mu_0^2}{2\pi \mu_0^2 + \omega^2} |\bar{v}_i|^2 d\omega \]
\[ = \frac{\theta_i^2 \mu_0^2}{2(\beta_i^2 - \mu_0^2)} \left( e^{-\mu_0 |\tau|} - e^{-\beta_i |\tau|} \right). \] This is again a symmetric function. It emerges whenever the covariance is taken between variables that are both affected by an Ornstein–Uhlenbeck process with the same delay characterized by rate \( \mu_0 \). We note that
\[ A_i(\tau) + A_i(-\tau) = 2S_i(\tau). \]
Auto-covariances and coefficients of variation

In terms of the functions $B(t)$, $A(t)$, and $S(t)$, the auto-

covariances for $\phi_i$, $\pi_i$, and $\mu$ are given by

$$
\frac{\chi_{\phi,\phi}(\tau)}{\phi_{0,i}} = (1 - 2\phi_{0,i})S_i(\tau) + \sum_j \phi_{0,j}^2 S_j(\tau),
$$

(51)

$$
\frac{\chi_{\pi,\pi}(\tau)}{\pi_{0,i}^2} = B_i(\tau) + 2C_i^\pi S_i(\tau) + \sum_j \left(C_j^\pi\right)^2 S_j(\tau),
$$

(52)

$$
\frac{\chi_{\mu,\mu}(\tau)}{\mu_0^2} = \sum_j \left[C_j^\mu \left(C_j^\mu + 2\phi_{0,j}\right)S_j(\tau) + \phi_{0,j}^2 B_j(\tau)\right].
$$

(53)

From these auto-covariances we obtain the CVs $\eta$:

$$
\eta^2_{\phi,i} = \left(\frac{\mu_0(1 - 2\phi_{0,i})}{2\beta_i(\beta_i + \mu_0)}\right) \theta_i^2 + \sum_j \left(\frac{\mu_0\phi_{0,j}}{2\beta_j(\beta_j + \mu_0)}\right) \theta_j^2
$$

(54)

$$
\eta^2_{\pi,i} = \left(\frac{1}{2\beta_i} + \frac{C_i^\pi\mu_0}{\beta_i(\beta_i + \mu_0)}\right) \theta_i^2 + \sum_j \left(\frac{2\beta_j(\beta_j + \mu_0)}{2\beta_j}\right) \theta_j^2
$$

(55)

$$
\eta^2_{\mu,i} = \sum_j \left(\frac{C_j^\mu(2\theta_j + 2\phi_{0,j}\mu_0)}{2\beta_j(\beta_j + \mu_0)} + \frac{\phi_{0,j}^2}{2\beta_j}\right) \theta_j^2
$$

(56)

Here we used equation (43) and the special cases $B_i(0) = \theta_i^2/(2\beta_i)$ and $A_i(0) = S_i(0) = \theta_i^2\mu_0 [2\beta_i(\beta_i + \mu_0)]^{-1}$ derived from equations (47), (48), and (49).

Intrinsic and extrinsic noise

Defining the shorthand

$$
S_i := S_i(0) = \frac{\theta_i^2\mu_0}{2\beta_i(\beta_i + \mu_0)} = \frac{\text{Var}(N_i)}{\phi_{0,i}^2 \text{Var}(\mu_0(\mu_0 + \beta_i))}
$$

(57)

as a compound parameter associated to noise source $i$, where $\text{Var}(N_i)$ is the variance of the Ornstein–Uhlenbeck process $N_i$, allows us to write equation (54) in the form of equation (14) of the main text, or as:

$$
\eta^2_{\phi,i} = (1 - \phi_{0,i})^2 S_i + \sum_{j \neq i} \phi_{0,j}^2 S_j,
$$

(58)

The two terms can be interpreted as intrinsic and extrinsic components of the noise. A discussion of this expression is given in the main text.

To see how the intrinsic and extrinsic components behave as a function of the protein’s mean mass fraction $\phi_{0,i}$, we assume that the relative noise amplitude $\theta_i$ scales as $\phi_{0,i}^{-1/2}$, so that $S_i \sim \phi_{0,i}^{-1}$. Equation (58) then predicts two regimes: at low expression levels, the intrinsic noise dominates and the CV scales as $\phi_{0,i}^{-1/2}$. At high expression levels, the intrinsic component becomes negligible compared to the extrinsic component, and the CV approaches a plateau. This type of scaling has indeed been observed in experiments [51] and has been predicted by other theory as well [16].

Cross-covariances and cross-correlations

The second half of the main text focuses on the analysis of cross-covariances between gene expression and growth. Here we give explicit expressions for the cross-covariances between expression levels $\phi_i$ protein production rates $\pi_i$, and the growth rate $\mu$; in conjunction with the CVs from equations (54), (55), and (56), these expressions can be used to compute the cross-correlation functions.

The cross-covariances are:

$$
\frac{\chi_{\phi,\mu}(\tau)}{\phi_{0,i} \mu_0} = C_i^\mu S_i(\tau) + \phi_{0,i} A_i(\tau)
$$

(59)

$$
\frac{\chi_{\pi,\mu}(\tau)}{\pi_{0,i} \mu_0} = C_i^\pi A_i(\tau) + \phi_{0,i} B_i(\tau)
$$

(60)

$$
\frac{\chi_{\phi,\phi}(\tau)}{\phi_{0,i} \phi_{0,j}} = -\phi_{0,i} S_i(\tau) - \phi_{0,j} S_j(\tau) + \sum_k \phi_{0,k}^2 S_k(\tau).
$$

(61)

$$
\frac{\chi_{\pi,\pi}(\tau)}{\pi_{0,i} \pi_{0,j}} = C_i^\pi A_i(\tau) + C_j^\pi A_j(\tau)
$$

(62)

In these expressions we have indicated the four noise modes as defined and illustrated in the main text.

The cross-covariance between expression levels of different proteins $\chi_{\phi,\phi}(\tau)$ is symmetrical. Any asymmetry that may be observed in this function will therefore remain unexplained by our model. Such an asymmetry could be explained by noise sources outside of protein production that act differently on the two proteins, or by gene regulation.

Whereas $\chi_{\phi,\phi}(\tau)$ is fully determined by the expression levels $\phi$ and the noise properties of the proteins, the cross-covariance between protein production rates $\chi_{\pi,\pi}(\tau)$ contains information about the GCCs. Measurements on this function may therefore shed light on the growth-control properties of proteins.
Results on growth-control coefficients

Relation between GCCs and FCCs

As stated in the main text, the growth-control coefficient $C'_{\mu}$, flux control coefficient $C'_{i}$, and mean mass fraction $\phi_{0,i}$ of a protein species $i$ are tightly related through

$$C'_{i} = C'_{\mu} + \phi_{0,i}. \quad (63)$$

Here, we provide a derivation. (See also [47, 56] for a similar derivation and application in an evolutionary context.)

By definition, $\mu_{d}(n) := J(n)/M(n)$; equation (63) then follows from the product rule for differentiation:

$$C'_{i} = \left[ \frac{\phi_{i}}{J(\phi)} \frac{\partial J(\phi)}{\partial \phi_{i}} \right]_{\phi_{0}}$$

$$= \frac{\phi_{0,i}}{\mu_{d}(\phi_{0})M(\phi_{0})} \left[ \frac{\partial [\mu_{d}(\phi)M(\phi)]}{\partial \phi_{i}} \right]_{\phi_{0}}$$

$$= \frac{\phi_{0,i}}{\mu_{d}(\phi_{0})} \left[ \frac{1}{M(\phi_{0})} \left( \frac{\partial \mu_{d}}{\partial \phi_{i}} \right)_{\phi_{0}} + \mu_{d}(\phi_{0}) \left( \frac{\partial M}{\partial \phi_{i}} \right)_{\phi_{0}} \right]$$

$$= C'_{i} + \phi_{0,i}. \quad (64)$$

In this derivation, we used that $\partial M(\phi)/\partial \phi_{i} = 1$ and $M(\phi_{0}) = 1$, which both follow from the definition $M(n) := \sum n_{i}$.

Sum rules

In this section, we provide derivations of the sum rules for flux control coefficients (FCCs) and growth-control coefficients (GCCs).

In metabolic control analysis [45, 46], the FCC of enzyme $i$ with respect to a metabolic flux $J$ has been defined as

$$C'_{i} := \left[ \frac{\phi_{i}}{J(\phi)} \frac{\partial J(\phi)}{\partial \phi_{i}} \right]_{\phi_{0}} \quad (65)$$

where $\phi_{i}$ is the concentration of enzyme $i$. A classical result is the sum rule for FCCs:

$$\sum_{i} C'_{i} = 1. \quad (66)$$

This result is ultimately based on the common assumption that metabolic fluxes are extensive quantities; that is, for any $\alpha > 0$ and any composition $\phi$ it is assumed that

$$J(\alpha \phi) = \alpha J(\phi). \quad (67)$$

Differentiating both sides of this equation with respect to $\alpha$ and evaluating the result in $\alpha = 1$ yields:

$$\sum_{i} \phi_{i} \frac{\partial J}{\partial \phi_{i}} = J(\phi), \quad (68)$$

where the summation runs over all protein species. Dividing by $J$ directly results in the sum rule of equation (65).

Our definition of growth-control coefficients (GCCs) proceeds analogously, as

$$C'_{\mu} := \left[ \frac{\phi_{i}}{\mu_{d}(\phi)} \frac{\partial \mu_{d}}{\partial \phi_{i}} \right]_{\phi_{0}} \quad (69)$$

For GCCs, however, a different sum rule holds:

$$\sum_{i} C'_{\mu} = 0. \quad (70)$$

The derivation of this result is directly analogous to that for FCCs, except that $\mu_{d}$ is assumed to be intensive rather than extensive. That is, for any $\alpha > 0$ and any composition $\phi$ it is assumed that

$$\mu_{d}(\alpha \phi) = \mu_{d}(\phi). \quad (71)$$

Differentiating both sides with respect to $\alpha$ and evaluating the resulting expression in $\alpha = 1$ now gives

$$\sum_{i} \phi_{i} \frac{\partial \mu_{d}}{\partial \phi_{i}} = 0. \quad (72)$$

Dividing by $\mu_{d}$ directly results in the sum rule of equation (70).
Application to the two-protein toy model

In the main text, a toy model is presented in which the cell contains only two protein species, called X and Y. We here derive the mathematical expressions underlying these results.

\[
R_{\phi_{\nu,\mu}}(\tau) = \frac{C^\mu_X S_Y(\tau) + \phi_{0,Y} A_Y(\tau) - \phi_{0,Y} \left[ C^\mu_X S_Y(\tau) + \phi_{0,Y} A_Y(\tau) \right] + \phi_{0,X} \left[ C^\mu_X S_X(\tau) + \phi_{0,X} A_X(\tau) \right]}{\eta_{\nu,\eta_{\mu}}}
\]

\[
R_{\tau_{\nu,\mu}}(\tau) = \frac{C^\mu_X A_Y(-\tau) + \phi_{0,Y} B_Y(\tau) + C^\mu_X S_X(\tau) + \phi_{0,X} A_X(\tau)}{\eta_{\tau_{\nu,\eta_{\mu}}}}
\]

where the coefficients of variation in the denominators are given by

\[
\eta^2_{\phi_{\nu}} = \left( \frac{1 - \phi_{0,Y}^2}{2 \beta_Y (\beta_Y + \mu_0)} + \frac{\phi_{0,Y}^2 \beta_Y + \mu_0}{2 \beta_Y (\beta_Y + \mu_0)} \right)
\]

\[
\eta^2_{\tau_{\nu}} = \left( 1 + \frac{\mu_0 C^\mu_X (C^\mu_X + 2)}{\beta_X + \mu_0} \right) \frac{\phi_{0,X}^2}{2 \beta_X}
\]

\[
\eta_{\mu} = \left( \frac{\mu_0 C^\mu_X (C^\mu_X + 2 \phi_{0,X})}{\beta_X + \mu_0} + \phi_{0,X}^2 \right) \frac{\phi_{0,Y}^2}{2 \beta_Y}.
\]

The expressions above can be further simplified by inserting the constraints

\[
C^\mu_X + C^\mu_Y = 0,
\]

\[
\phi_{0,X} + \phi_{0,Y} = 1.
\]

Regulation in the two-protein model

In deriving the results above, we assumed that \( f \) is constant; that is, the synthesis rate of the proteins is not dynamically regulated in response to fluctuations in the protein mass fractions. Generally, if gene regulation is included, analytical results become hard to obtain. The simple case of the two-protein toy model is an exception, as we now show.

Assuming, as before, that \( f \) is an intensive quantity, we find the additional constraint (sum rule)

\[
C^f_X + C^f_Y = 0.
\]

Cross-correlations

For the two-protein model, the general expressions of equations (45), (59) and (60) evaluate to the following cross-correlations \( R_{\phi_{\nu,\mu}}(\tau) \) and \( R_{\tau_{\nu,\mu}}(\tau) \):

\[
\text{Control Autogenic} \quad \text{Dilution}
\]

\[
\text{Control Autogenic} \quad \text{Transmission}
\]

Furthermore, \( f \) denotes the fractional allocation of resources towards each protein species, so that

\[
f_X + f_Y = 1.
\]

Thereby, the regulation has only one degree of freedom and is thus fully characterized by just one parameter:

\[
C^{f_Y}_X = C^{f_Y}_Y = -C^{f_X}_X = -C^{f_X}_Y,
\]

which quantifies the auto-regulation of protein X, that of Y, the regulation of Y by X, and vice versa.

These constraints allow us to calculate the two-protein cross-correlations analytically. Equation (35) simplifies to

\[
\frac{\delta \phi_Y}{\phi_{0,Y} \mu_0} = C^f_X \frac{\delta \phi_X}{\phi_{0,X}} + C^f_Y \frac{\delta \phi_Y}{\phi_{0,Y}}
\]

\[
+ (1 - \phi_{0,Y} (\nu_Y - \nu_X)) - \frac{\delta \phi_Y}{\phi_{0,Y}}
\]

\[
= -C^{f_Y}_X \frac{\delta \phi_X}{1 - \phi_{0,Y} + \mu_0} + C^{f_Y}_Y \frac{\delta \phi_Y}{\phi_{0,Y}}
\]

\[
+ (1 - \phi_{0,Y} (\nu_Y - \nu_X)) - \frac{\delta \phi_Y}{\phi_{0,Y}}
\]

\[
= \left( \frac{C^{f_Y}_Y}{1 - \phi_{0,Y}} - 1 \right) \frac{\delta \phi_Y}{\phi_{0,Y}}
\]

\[
+ (1 - \phi_{0,Y} (\nu_Y - \nu_X)).
\]

The Fourier transform then reads

\[
\overline{\delta \phi_Y} = (1 - \phi_{0,Y}) \left( \frac{\mu_0}{\mu_Y + \overline{\omega}} \right) (\overline{\nu_Y} - \overline{\nu_X}).
\]

This equation differs only slightly from equation (36): apart from a constant prefactor \( 1 - \phi_{0,Y} \), a term \( \mu_0 \) in the denominator is replaced by

\[
\mu_Y = \mu_0 \left( 1 - \frac{C^{f_Y}_Y}{1 - \phi_{0,Y}} \right).
\]
Because equations (36) and (84) have the same form, their solutions are analogous.

To illustrate the effect of regulation in this toy model, the expression–growth cross-correlation function is plotted in Fig S2A, for different values of $C_Y^0$. Positive auto-regulation widens the cross-correlation function, indicating increased auto-correlation timescales; negative auto-regulation narrows it, indicating decreased auto-correlation timescales. Without regulation, dilution by growth quenches fluctuations in the concentration $\phi_Y$ with the associated time scale $\mu_Y^{-1}$. In the presence of regulation this time scale is adjusted to $\mu_Y^{-1}$. This directly determines the time scales of the cross-correlations.

Further, regulation affects the CV of the expression of both X and Y, because the term $\mu_0$ in the denominator of equation (75) is replaced by $\mu_Y$. This is shown in Fig S2B. If $C_Y^0$ is negative (negative auto-regulation), fluctuations are actively quenched; the variance of $\phi_Y$ is therefore reduced. Conversely, if $C_Y^0$ is positive (positive auto-regulation) fluctuations are amplified; the variance in $\phi_Y$ is therefore increased.

**Stochastic simulation of the two protein model**

To test the analytical results on the two-protein toy model, we used the Euler–Maruyama method to propagate the system of stochastic differential equations and thus calculated the cross-correlations numerically. These simulations are based on the non-linear equations (4), (13) and (14), and a non-linear function $\mu_i(\phi)$ was used. As such, they demonstrate that, provided $\mu_i(\phi)$ is sufficiently smooth, the approximations used in the derivation of the linearized equations are appropriate. The results are included in Fig 3AB.

In this simulation we chose the parameters of the non-linear system such as to match the parameters used for the plots in Fig 3AB. Specifically, we specified the allocation coefficient as $f_Y = 0.33$, and the parameters of the noise sources as $\theta_Y = 0.5$, $\theta_X = 0.5$, $\beta_Y = 4/h$, and $\beta_X = 4/h$, as in Fig 3AB. For the the deterministic part of the growth rate we chose $\mu_i(\phi) := (\alpha_X \phi_X + \alpha_Y \phi_Y)/(\phi_X + \phi_Y)$, with $\alpha_X = 0.63/h$ and $\alpha_Y = 1.37/h$, such that $\mu_0 = \mu_i(\phi_0) \approx \mu_i(f) = 1/h$ and the control coefficients would approximate the ones in Fig 3 ($C_Y^0 = -C_X^0 = 0.25$).

We propagated the dynamical equations in discrete time steps $\Delta t$ by repeating the following calculations. (Here, the index $i$ is either X or Y.)

\[
\begin{align*}
\mu_i(t) &= (\alpha_X \phi_X(t) + \alpha_Y \phi_Y(t))/(\phi_X(t) + \phi_Y(t)), \\
\varpi_i(t) &= f_i^2 \mu_i(t) + N_i(t), \\
\mu_i(t + \Delta t) &= \mu_i(t) + N_i(t + \Delta t), \\
\phi_i(t + \Delta t) &= \phi_i(t) (1 - \Delta \mu_i(t)) + \Delta \varpi_i(t), \\
N_i(t + \Delta t) &= N_i(t)(1 - \beta_i \Delta t) + \sqrt{\beta_i} \Delta t dW(t).
\end{align*}
\]

In the last line, $dW(t)$ is random variable drawn from a standard normal distribution each time step. We ran the simulation for roughly 30,000 doubling times, using time steps $\Delta t = 10^{-3}$.

After the simulation, we measured $\phi_Y$ to test whether $f_Y \approx \phi_Y$ as is assumed in section 2.5. Indeed, $f_Y$ did not differ from $\phi_Y$ by more than 5%. We next verified that $\mu_0 := \mu_i(\phi_0) \approx 0.993/h$ was close to $1/h$ and that $C_Y^0 = \phi_0 \varpi_i(\alpha_Y/\mu_0 - 1) \approx 0.249$ was close to 0.25, as intended.

**Application to the many-protein model**

As described in the main text, we applied the modeling framework to model a full cell, containing 1021 protein species in total. We here provide the mathematical details of this model. First we explain how the operon construct was represented. Then we describe our parameter choices and how they were informed by previously published experimental data.

**GFP reporter fused to the lac operon**

As mentioned in the main text, Kiviet et al. performed experiments in which the expression levels of the genes lacY and lacZ were reported by a green fluorescent protein (GFP), whose gene was transcriptionally fused to the lac operon [5]. Thus, the GFP signal was used as a proxy for the concentration of the lac proteins LacY and LacZ. To include this construct in the model, we assumed that the synthesis noise of each protein, introduced by transcription of the operon and the translation of its genes, can be linearly decomposed into two components. The first component reflects noise originating from processes that similarly affect all genes on the operon, adding to the correlation between their expression levels. In particular, this includes fluctuations in the copy number of the mRNA transcript due to transcriptional noise. The second component reflects noise originating from processes that affect each gene independently, including all post-transcriptional processes.

Formally, we included noise sources $N_Y$, $N_Z$, and $N_O$, respectively acting on LacY, LacZ, and GFP synthesis only, as well an additional noise source $N_O$ that simultaneously affects the synthesis of all of these proteins. That is, equation (13) was modified to:

\[ \pi_i = f_i \mu_i(\phi_0) + N_i + \iota_i^O \frac{\pi_{0,i}}{\pi_{0,O}} N_O. \]  

Here, $\iota_i^O$ is an indicator function that specifies whether protein $i$ is encoded on the lac operon: $\iota_i^O = 1$ for $i \in \{Y, Z, G\}$ and $\iota_i^O = 0$ otherwise.

As before, all noise sources are mutually independent Ornstein–Uhlenbeck processes with zero mean. For simplicity, we incorporated the operon noise such that it affects the three genes in proportion with their average production rate. In the limit that $\nu_i = 0$ for $i \in O$ and $N_O \neq 0$, the three proteins behave as a single protein O, with the same scaling $N_j = \pi_{0,j} \nu_j$ as the other genes.
After linearization (cf. equation (25)) we arrive at
\[
\frac{\delta \pi_i}{\pi_{0,i}} = \sum_j C_{i,j}^\mu \frac{\delta \phi_j}{\phi_{0,j}} + \nu_i + I_i^\mu \nu_0. \tag{88}
\]
Here, the summation runs over all 1021 proteins, including Y, Z, and G. Moreover, we have defined \( \nu_0 := \frac{N_0}{\pi_{0,0}} = \frac{N_0}{\pi_{0,Y} + \pi_{0,Z} + \pi_{0,G}} \). To obtain the cross-covariances, the analysis now proceeds exactly as before. Using the shorthand \( \phi_{0,i} := \phi_{0,Y} + \phi_{0,Z} + \phi_{0,G} \) and \( C_\mu^O := C_\mu^Y + C_\mu^Z + C_\mu^G \), the resulting (cross-)covariances are given by

\[
\begin{align*}
X_{\phi,\phi}(\tau) &= \frac{(1 - 2\phi_{0,i})S_i(\tau) + I_i^\mu (1 - 2\phi_{0,O})S_O(\tau) + \sum_j \phi_{0,j}^2 S_j(\tau)}{\pi_{0,i}^2}; \\
X_{\pi,\pi}(\tau) &= \frac{B_i(\tau) + 2C_i^\mu S_i(\tau) + I_i^\mu \left[ B_O(\tau) + 2C_O^\mu S_O(\tau) \right] + \sum_j (C_j^\mu)^2 S_j(\tau)}{\mu_0^2}; \\
X_{\mu,\mu}(\tau) &= \frac{\sum_j \left\{ \phi_{0,j}^2 B_j(\tau) + \left( C_j^\mu \right)^2 \phi_{0,j}^2 S_j(\tau) \right\}}{\mu_0^2}; \\
X_{\pi,\mu}(\tau) &= \frac{\phi_{0,i} B_i(\tau) + C_i^\mu S_i(\tau) - \sum_j \phi_{0,j} \left[ C_j^\mu S_j(\tau) + \phi_{0,j} A_j(\tau) \right]}{\pi_{0,i} \mu_0}; \\
X_{\mu,\pi}(\tau) &= \frac{\phi_{0,i} B_i(\tau) + C_i^\mu S_i(\tau) - \sum_j \phi_{0,j} \left[ C_j^\mu S_j(\tau) + \phi_{0,j} A_j(\tau) \right]}{\pi_{0,i} \mu_0}; \\
\end{align*}
\]

The index \( j \) of the summations now runs over all proteins as well as the operon \( j = O \), so that all noise sources are included.

### Parametrization

In order to simulate a whole cell, we had to specify four parameters for each protein \( i \), namely its abundance \( \phi_{0,i} \), its control coefficient \( C_i^\mu \), and its noise properties \( \theta_i \) and \( \beta_i \). In addition, the mean growth rate \( \mu_0 \) had to be set for each of the three growth states presented in the main text (low, intermediate, and high).

### Monod curve

Under the experimental conditions, the mean growth rate \( \mu_0 \) follows an empirical Monod curve \( \mu_M(\phi_{0,O}) \) as a function of the mean expression level \( \phi_{0,O} \) of the lac operon [5]. To ensure that our parameters respect this relation, we constrain them with the Monod curve shown in Fig 4B,

\[
\mu_M(\phi_{0,O}) := \mu_{\text{max}} \frac{\phi_{0,O}}{\phi_{0,O} + \phi_{\text{half}}}, \tag{95}
\]

which is defined by a maximal growth rate of \( \mu_{\text{max}} = 0.8 \) h\(^{-1}\) and \( \phi_{\text{half}} = 0.005 \). The parameter \( \phi_{\text{half}} \) was taken clearly larger than the abundance of an average protein, since transporter proteins are highly expressed in growth media where they are utilized [29].

### Choosing the average growth rate \( \mu_0 \)

The average growth rates in the three growth states were chosen in rough agreement with the experiments as \( \mu_0 = 0.19, 0.45, \) and \( 0.75 \) h\(^{-1}\).

### Choosing reversion rates \( \beta \)

For simplicity, we assumed that the reversion rates \( \beta_i \) of all noise sources were identical and considerably larger than the average growth rate \( \mu_0 \); that is, we set \( \beta_i = \beta := 4\mu_{\text{max}} \) for all \( i \).

### Choosing \( \phi_0 \) and \( \theta \) based on experimental data

Several experimental datasets are available that provide cell-wide estimates of \( E. \colli \) protein abundances [28, 29, 51, 53]. In particular, Taniguchi et al. have measured the means and variances of 1018 protein abundances [51]. We used these measurements to fix the values of \( \phi_{0,i} \) and \( \theta_i \) for all proteins \( i \) in our model, as we now explain.

First we determined the average mass fractions \( \phi_0 \). Given the mean growth rate \( \mu_0 \) for a given growth condition, the average mass fraction of the operon, \( \phi_{0,O} \), follows directly from the Monod curve described above. We then assumed
equal properties for the proteins on the operon, i.e. \( \phi_{0G} = \phi_{0Y} = \phi_{0Z} = \phi_{0O}/3 \) and \( \theta_{0G} = \theta_{0Y} = \theta_{0Z} \). We denote the sum of all measured protein abundances as \( M \), and the mass of the whole cell including the lac operon as \( M \). Then \( M = m/(1 - \phi_{0O}) \), and the average mass fraction of protein \( i \) can be calculated by dividing the measured abundance by \( M \).

Next, we calculated the values of the noise amplitudes \( \phi \). Equation (92) shows that the squared coefficient of variation of any given protein species is a linear function of the squared noise amplitudes of all noise sources. Conversely, imposing the measured variances of all protein species results in a linear system of equations, the solution of which uniquely determines the \( \theta_i^2 \) (for given \( \mu_0, \phi_0 \) and \( \beta \)).

To do so explicitly, we assigned all proteins an index in the range 1 to \( K + 3 \), where \( K \) is the number of proteins that are not encoded on the lac operon (in this case \( K = 1018 \)); proteins G, Y and Z were assigned indices 1 to 3, respectively. We set the amplitude of the operon noise source to \( \theta_0 = a\theta_G = a\theta_Y = a\theta_Z \), where \( a \) defines the ratio of the amplitude of \( \theta_0 \) to the amplitudes of \( \theta_G, \theta_Y \), and \( \theta_Z \). We used \( a = 1.5 \) to obtain the results in the main text (Figure 4). This procedure allowed us to rewrite equation (92) in \(( (K + 3) \times (K + 3) \) dimensional) matrix form:

\[
\eta_0^2 = T \vartheta^2, \tag{96}
\]

where the vector \( \vartheta^2 \) contains the squared noise amplitudes, \( \eta_0^2 \) is the vector of squared CVs, and the elements of matrix \( T \) are given by

\[
T_{ij} = \frac{\mu_0}{2\beta(\beta + \mu_0)} \times \begin{cases} 
(1 - \phi_{0G})^2 & \text{if } i = j = 1, \\
(1 - \phi_{0G})^2 + (1 - \phi_{0O})^2 & \text{if } i = j \neq 1, \leq 3, \\
\phi_{0G}^2 + (1 - \phi_{0O})^2 & \text{if } j = 1, i > 3, \\
\phi_{0G}^2 + \phi_{0O}^2 & \text{if } i = j > 3, \\
(1 - \phi_{0i})^2 & \text{if } i \neq i, j, \\
\phi_{0j}^2 & \text{if } i \neq i, j, \\
(1 - \phi_{0j})^2 & \text{if } i = j \neq 1, \\
(1 - \phi_{0i})^2 & \text{if } i = j \neq 1, \\
\end{cases} \tag{97}
\]

Lastly, \( \vartheta^2 \) was obtained by numerically inverting equation (96).

### Choosing the growth-control coefficients

Large-scale measurements of growth-control coefficients are not readily available. Therefore, we randomly sampled the GCCs using the following heuristic. First, we randomly assigned proteins from the database measured by Taniguchi et al. to the non-metabolic H sector, until the total H sector size was \( \sim 25\% \) of the sum of measured protein abundances \( m \). The size of this non-metabolic mass fraction has been directly estimated as 25% by O’Brien et al. [48]; it has been described earlier by Hui et al. as part of a growth-rate independent proteome fraction that amounts in total to about 60% of the total protein mass [28]. For these H proteins, we set \( C_{i}^\alpha = -\phi_i \) (note that \( \phi_i \) partly depends on the expression level of the lac operon, as described earlier), in line with equation (11) from the main text.

Second, the GCCs of the lac-operon proteins Y, Z, and G were chosen in agreement with the Monod curve discussed above. We next used the derivative of the Monod curve \( \mu_M(\phi_{0O}) \) to estimate \( C_0^\alpha \):

\[
C_0^\alpha := \left[ \frac{\phi_{0O}}{\mu_M} \frac{\partial \mu_M}{\partial \phi_{0O}} \right] \approx \frac{\phi_{0O}}{\mu_M} \frac{d\mu_M}{d\phi_{0O}} = \frac{\phi_{0O} - \phi_{half}}{\phi_{0O} + \phi_{half}}. \tag{98}
\]

Here we implicitly assumed that the operon as a whole is never so strongly over-expressed as to have a negative control on the growth rate.

Lastly, we chose \( C_Y^\alpha = C_Z^\alpha = (C_0^\alpha - C_G^\alpha)/2 \) for simplicity.

Third, for the remaining proteins (which are neither in the H sector nor encoded on the lac operon), we assumed that the GCCs tend to scale with their abundance, and enforced the sum rule. In practice, we drew a uniformly distributed random number \( x_i \in [0, 1] \) for each protein that was not in the H sector and set

\[
C_i^\alpha = \frac{x_i\phi_0}{\sum_{j \in H} x_j \phi_0} \left( \phi_H - C_0^\alpha \right). \tag{99}
\]

As in the main text, \( \phi_H \) is the mass fraction of all H proteins combined. To determine the robustness of our results with respect to the sampling scheme, we also tried different GCC distributions, for example using exponentially distributed variables \( x_i \). This yielded visually indistinguishable results.

### Alternative choice for abundances and variances

Although the data set obtained by Taniguchi et al. [51] is unique to report both abundances and variances for each measured protein, it suffers from a systematic underestimation of the protein masses [29, 53]. We therefore also estimated the protein abundances from another proteomics dataset, using a phenomenological model to calculate the variances in the absence of explicit measurements.

In this case, the abundances \( \phi_{0i} \) were taken from a dataset obtained from Arike et al. [53]. Given that the mean copy number of protein \( i \) is \( n_i \), a theoretical minimum can be calculated for its variance; it is given by \( n_i^2 \sigma_{ab} + n_i \beta_W \), with \( \sigma_{ab} = 0.025 \) and \( \beta_W = 450 \) [68]. The original paper reported variances that were typically higher than this lower limit [68]. To empirically match these results, we chose

\[
\text{Var}_i = n_i^2 \sigma_{ab} + n_i(\beta_W + x_i). \tag{100}
\]

with \( x_i \) an exponentially distributed random variable with mean 200.

The resulting cross-correlations are plotted in Fig S3. They lead to the same qualitative conclusions as the analysis based on the data from Taniguchi et al.
Alternative model: Noise in the allocation of the flux

Throughout this document we have considered noise sources that act on each protein-synthesis rate independently. Alternatively, one could hypothesize that the observed fluctuations in protein concentrations instead originate from noise on the allocation of the available metabolic flux. Especially when translation is a highly rate-limiting step in protein production, so that mrnas compete for a limited capacity of ribosomes, a stochastic increase in the synthesis of a given protein species may necessarily go at the expense of the synthesis rates of other protein species. In this section, we therefore explore a model in which such “allocation noise” is dominant instead of the pure “production noise” studied above.

In the modified model, each protein species is allotted a fluctuating fraction $f_i(t)$ of the total metabolic flux, with the additional constraint that $\sum_i f_i(t) = 1$ at all times. We end up with an adjusted version of equation (13):

$$\pi_i := f_i(t) \mu_d(\phi)$$

with

$$f_i(t) = \frac{f_{0i} + N_i}{\sum_j f_{0j} + N_j} = \frac{f_{0i} + N_i}{1 + \sum_j N_j}.$$  

Here, the $N_i$ are again independent Ornstein–Uhlenbeck processes. This also results in an adjusted version of equation (14):

$$\mu = \mu_d(\phi).$$

As before, in the small-noise limit $\mu_d$ can be linearized around the mean composition $\phi_0$:

$$\mu_d(\phi) := \mu_0 \left( 1 + \sum C_i^\phi \left( \frac{\phi_i}{f_{0i}} - 1 \right) \right),$$

where we assumed that $\phi_{0i} \approx f_{0i}$, as previously.

Unfortunately, production rates as defined in equation (101) are intrinsically correlated and therefore analytical expressions can not be obtained within our current framework. To gain insight in the cross-correlations generated by this system, we therefore numerically integrated equations (101) and (104) for a model cell containing 40 proteins with randomly sampled parameters (see next section for details about parameters and the simulation method). The cross-correlations obtained for one of the proteins (protein 1) are shown in Fig S5A.

A striking observation is that the $\phi - \mu$ cross-correlations of the model with “allocation noise” (Fig S5A and C) are always perfectly symmetrical, with their peak at zero delay ($\tau = 0$). In itself, the presence of a symmetric mode is no surprise. Indeed, an increase in a protein species’ synthesis rate, affects—with a delay—its concentration, and this in turn affects the growth rate, in proportion with the protein’s GCC. Therefore, a symmetrical control mode is to be expected.

However, when noise acts dominantly on the allocation of the flux, the growth rate is not instantly affected by fluctuating production rates. Therefore, there is now pathway by which noise in a protein’s production rate instantly affects the growth rate, and then—afer a delay—the protein concentrations (either by dilution or by direct production). This removes the asymmetric autogenic mode that caused the asymmetry in the $\phi - \mu$ cross-correlation as well as the symmetric mode in the $\phi - \mu$ cross-correlation in the “production noise” model (see also Fig 3C–F of the main text).

To check the the symmetric $\phi - \mu$ cross-correlation is indeed a typical feature of the allocation noise model and not just a feature of the particular cell sampled, we analytically calculated cross-correlations for a similar cell (i.e. same protein variances, same mean protein abundances, same GGCs), but now according to the “production noise” model (i.e. production rates according to equation (13)). The resulting cross-correlations are shown in Fig S5B and show clear asymmetry in the $\phi - \mu$ cross-correlation (demonstrated in Fig S5C).

Since the measured cross-correlations by Kiviet et al. showed clear asymmetrical features as well [5]—which in fact prompted them to propose a “common-noise” source—we conclude that noise on allocation alone cannot explain the observed cross-correlations.

Simulations of the “allocation-noise” model

We here provide details on the simulations discussed in the previous section and presented in Fig S5.

We simulated a cell with 40 proteins. To choose their mean protein fractions, we drew uniformly distributed numbers ($x_j \sim \mathcal{U}[0, 1]$, where $\mathcal{U}[0, 1]$ is the uniform distribution over the interval $[0, 1]$) and normalized them to obtain protein fractions, i.e., $\phi_{i,0} = f_{i,0} = x_i / \sum_j x_j$. Approximately 60% of the proteome was assumed to be metabolic. For the $\approx 40\%$ non-metabolic proteins, we set $C_i^\phi = -\theta_{i,0}$ for the remaining proteins we sampled GCCs according to a uniform distribution, scaled with the protein’s mean fraction, under the constraint of the sum rule. That is: for any metabolic protein $i$, draw an $x_i \sim \mathcal{U}[0, 1]$, and subsequently set $C_i^\phi = (\sum_{j \in H} \phi_j) (x_i \phi_i) / (\sum_{j \in H} x_j \phi_j)$. We also sampled the noise amplitudes $\theta_{AN}$ (here the label “AN” stands for “allocation noise”) uniformly and then scaled them with $\sqrt{f_{0i}}$, i.e., $\theta_{AN,j} \sim \left[0, 1 / \sqrt{f_{0j}} \right] \mathcal{U}[0, 1]$.

First we numerically integrated Eq.(101)-(104) using the numerical integration scheme below to calculate the cross-correlation for the allocation-noise model.

$$\phi(t + \Delta t) = \phi(t) \left( 1 - \Delta t \sum_i \pi_i(t) \right) + \Delta t \pi(t),$$

$$N(t + \Delta t) = N(t) (1 - \Delta t \beta) + \sqrt{\Delta t} \theta_{AN} * dW(t).$$
\[ \pi(t + \Delta t) = \mu_0 \left( \frac{f_0 + N(t + \Delta t)}{1 + \sum_i N_i(t + \Delta t)} \right) \left( 1 + \sum_i C_i^\mu \left( \frac{\phi_i(t + \Delta t)}{f_{0,i}} - 1 \right) \right)^\text{noise} \]

(107)

Here, \( \ast \) denotes element-wise multiplication.

To define a similar cell in the “production-noise” model, we first measured \( \eta_{\phi,AN} \), in the “allocation noise” simulation and used a method similar to equation (96) to calculate which noise amplitudes \( \theta_{PN} \) (PN stands for “production noise”) would create the same variances in the “production-noise” model:

\[ \vartheta_{PN}^2 = (T_{PN})^{-1} \eta_{\phi,AN}^2 \]

With the matrix \( T_{PN} \) similar to equation (97):

\[ (T_{PN})_{ij} = \frac{\mu_0}{2\beta(\beta + \mu_0)} \left\{ \begin{array}{ll} (1 - f_{i,0})^2, & \text{if } i = j, \\ f_{i,j}^2, & \text{else.} \end{array} \right. \]