Supplementary Information

The Supplementary Information derives Eq. (1) in the main text and proves that the extrinsic term indeed can be interpreted as the normalised covariance between two identical and independent systems embedded in the same environment. It also briefly discusses transcription noise, plasmid extinctions and how well GFP reports plasmid copy numbers.

Deriving Eq. (1)

Process and method When the birth and death process behind Eq. (1) in the main text is in state $\{n_1, n_2\}$, the possible transitions are

$$\begin{array}{ccccc} & & \{n_1, n_2+1\} & & \\ & & \uparrow\ R_2^+ & & \\ \{n_1-1, n_2\} & \xleftarrow{R_1^-} & \{n_1, n_2\} & \xrightarrow{R_1^+} & \{n_1+1, n_2\} \\ & & \downarrow\ R_2^- & & \\ & & \{n_1, n_2-1\} & & \end{array} \tag{S1}$$

where R_1^- and R_1^+ depend on n_1, and R_2^- and R_2^+ depend on both n_1 and n_2. The corresponding Markov process is governed by master equation

$$\begin{aligned} \frac{\partial}{\partial t} p_{n_1 n_2} = {} & R_1^+(n_1-1)\, p_{n_1-1 n_2} + R_1^-(n_1+1)\, p_{n_1+1 n_2} + R_2^+(n_1, n_2-1)\, p_{n_1 n_2-1} + R_2^-(n_1, n_2+1)\, p_{n_1 n_2+1} \\ & -\left[R_1^+(n_1) + R_1^-(n_1) + R_2^+(n_1, n_2) + R_2^-(n_1, n_2)\right] p_{n_1 n_2} \end{aligned} \tag{S2}$$

where $p_{n_1 n_2}$ is the joint probability of n_1 and n_2. When R_1 or R_2 is nonlinear in n_1 or n_2, Eq. (S2) must be approximated numerically or by expansion methods. A systematic analytical approach is to use van Kampen's Ω-expansion[1–3] that formulates all rates in terms of the system volume Ω such that discrete events give relatively small jumps when Ω is large. The lowest order expansion reproduces the macroscopic rate equations

$$\begin{aligned} \partial x_1 / \partial t &= r_1^+(x_1) - r_1^-(x_1) \\ \partial x_2 / \partial t &= r_2^+(x_1, x_2) - r_2^-(x_1, x_2) \end{aligned} \tag{S3}$$

where $x_i = n_i/\Omega$ and $r_i = R_i/\Omega$ in the limit $\Omega \to \infty$. Upper-case X thus refers to the molecular species, lower-case x to macroscopic concentrations, and the integer n to the number of molecules. The next order terms reproduce a version of the fluctuation-dissipation theorem (FDT), where stationary fluctuations follow

$$A\sigma + \sigma A^T + \Omega B = 0\,, \tag{S4}$$

where σ is the sought covariance matrix, A is the Jacobian of the macroscopic system and ΩB is a diffusion matrix depending on system size, stoichiometry and macroscopic reaction rates, all evaluated at steady state (see e.g. chapter 1.8 in Keizer's *Statistical Thermodynamics of Nonequilibrium Processes*[3] for an introduction, or Eqs. 3.46 and 6.115 in Risken's *The Fokker-Planck Equation*[4]). The reasons to invoke the Ω-expansion as a motivation for Eq. (S4) are that it provides a conceptually sound derivation, and that it ensures that A and ΩB in Eq. (S4) are defined macroscopically. The mesoscopic rates R_1 and R_2 should thus first be formulated in terms of Ω, so that the macroscopic rates r_1 and r_2 can be explicitly calculated, and then used to calculate $H_{ij} = \partial \ln(r_i^-/r_i^+)/\partial \ln(x_j)$ and other parameters. The examples of the main text should thus have referred to the macroscopic rates r rather than the mesoscopic R, which is why they do not explicitly include Ω. This informal use is motivated by the reduced notation, but the full notations (both mesoscopic rate equations and macroscopic master equations) should be used wherever space limitations allow.

Throughout the analysis, all parameters are evaluated at steady state. Many reactions are also assumed to be elementary complex[3] so that rapid transitions between states can be compounded into effective rates.

More intuitive parameters My strategy here is to rewrite A and ΩB in terms of averages of numbers of molecules, lifetimes and logarithmic gains. This is done purely by interpretations, never by restrictions or *ad hoc* assumptions. The Jacobian A can be written as

$$A_{ij} = \frac{\partial}{\partial x_j}\left(\frac{\partial x_i}{\partial t}\right) = \frac{r_i^+}{x_j}\frac{\partial \ln r_i^+}{\partial \ln x_j} - \frac{r_i^-}{x_j}\frac{\partial \ln r_i^-}{\partial \ln x_j} = -\frac{r_i}{x_j}\frac{\partial \ln\left(r_i^-/r_i^+\right)}{\partial \ln x_j} \equiv -\frac{r_i}{x_j} H_{ij} \tag{S5}$$

Parameter H_{ij} summarizes, in a scale-free way, how a change in x_j affects the balance between synthesis and degradation of x_i (see main text).

In nonlinear systems, the average number of molecules can be arbitrarily displaced from the point where the birth and death rates balance each other. However, within the approximation of Eq. (S4), it is definitionally true that $\langle n_i \rangle = \Omega x_i$. It may appear strange to use $n_i = \Omega x_i$ in the definition of the process but $\langle n_i \rangle = \Omega x_i$ in the interpretation of the result. But by the nature of the approach, A and ΩB are defined macroscopically where n_i displays zero relative deviations from $\langle n_i \rangle$.

The average lifetime can in turn be defined by the concentration divided by the total rate of formation or elimination, $\tau_i = x_i / r_i$. For exponential decay, $r_i^- = d_i x_i$ this trivially gives $\tau_i = 1/d_i$, but for nonlinear elimination mechanisms, average lifetimes depend on concentrations. If the concentrations change due to fluctuations or adjustments, it is difficult to calculate the real average lifetimes since the rate of degradation changes. However, that is not a problem here as A and ΩB are evaluated at an asymptotically stable macroscopic steady state where all x_i / r_i are constant. Each molecule is then surrounded by a constant environment of other molecules, and the lifetime definitions are both exact and universal, resulting in

$$\Omega r_i = \langle n_i \rangle / \tau_i \,. \tag{S6}$$

Even if the main text only considers examples with exponential decay, Eq. (1) is thus more general than that, and Eq. (S5) can always be rewritten as

$$A_{ij} = -\frac{x_i}{x_j}\frac{r_i}{x_i} H_{ij} = -\frac{\langle n_i \rangle}{\langle n_j \rangle}\frac{H_{ij}}{\tau_i} \tag{S7}$$

where angular brackets denote averages. For the particular process considered here, the diffusion matrix B similarly follows[4]

$$\Omega B_{ii} = 2\Omega r_i = 2\langle n_i \rangle / \tau_i \quad \text{and} \quad \Omega B_{ij} = 0 \ \text{ for } i \neq j \tag{S8}$$

where the 2:s reflect that births and deaths add and eliminate one molecule each, and the 0:s reflect that n_1 and n_2 are produced and consumed in separate reactions.

Solving the equation system. To facilitate calculations and interpretations I introduce the normalized matrices

$$\begin{aligned} V_{ij} &= \sigma_{ij} / (\langle n_i \rangle \langle n_j \rangle) = V_{ji} \\ M_{ij} &= A_{ij} \langle n_j \rangle / \langle n_i \rangle \\ D_{ij} &= \Omega B_{ij} / (\langle n_i \rangle \langle n_j \rangle) = D_{ji} \end{aligned} \,, \tag{S9}$$

This changes Eq. (S4) to

$$MV + (MV)^T + D = 0 \,. \tag{S10}$$

which could be further divided into separate matrices containing averages of numbers of molecules, logarithmic factors, lifetimes and stoichiometric factors. Eqs. (S1)-(S3) give

$$M = -\begin{bmatrix} \dfrac{H_{11}}{\tau_1} & 0 \\ \dfrac{H_{21}}{\tau_2} & \dfrac{H_{22}}{\tau_2} \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} \dfrac{2}{\tau_1 \langle n_1 \rangle} & 0 \\ 0 & \dfrac{2}{\tau_2 \langle n_2 \rangle} \end{bmatrix}. \tag{S11}$$

The only thing left is to solve the corresponding Eq. (S10) for V using standard methods for linear equation systems. When the dust has settled, this gives

$$V_{11} = \frac{1}{\langle n_1 \rangle H_{11}} \tag{S12}$$

$$V_{12} = -V_{11}\frac{H_{21}}{H_{22}}\frac{H_{22}/\tau_2}{H_{11}/\tau_1 + H_{22}/\tau_2} = V_{21} \,, \tag{S13}$$

$$V_{22} = \frac{1}{\langle n_2 \rangle H_{22}} + V_{11}\frac{H_{21}^2}{H_{22}^2}\frac{H_{22}/\tau_2}{H_{22}/\tau_2 + H_{11}/\tau_2} \tag{S14}$$

as in Eq. (1). The reason for formulating the solution in these terms can be understood by looking at the different limits where: n_2 deterministically follows n_1; n_1 does not fluctuate; n_1 stays at fixed random values; $\langle n_2 \rangle$ is proportional to $\langle n_1 \rangle$;

and n_1 fluctuates very rapidly. Each limit eliminates entire terms or factors. Numbers n_1 and n_2 are negatively correlated when H_{21} is positive since an increase in n_1 then causes an average decrease in n_2. Autocorrelations follow directly from Onsager's regression hypothesis[4], involving the exponential of the Jacobian.

Covariances in the Two-gene Study

The second result that remains to be proven is that the extrinsic noise term in Eq. (1) can be interpreted as the normalized covariance between two identical and independent systems X_2 and X_3 in the same environment X_1. The theoretical study[5] that motivated the two-gene experimental strategy[6] used a general argument based on an integral over all intrinsic and extrinsic variables in the system. However, that formulation implicitly assumes that the intrinsic and extrinsic variables stay at fixed random values, i.e., they were not defined by random processes but by static associations between variables. It does not cover the present scenario where n_1 and n_2 change randomly over time. Here I show that the experimental strategy is indeed supported by the model above. This requires an extension to three dimensions, with H_{ij}=0 for ij = [12, 32, 13, 23]. The zeros in the Jacobian reflect the assumption of disorder and independence, while the assumption of equivalence adds the simplifications $M_{21} = M_{31}$, $M_{22} = M_{33}$ and $D_{22} = D_{33}$. The statement we wish to prove is that the normalized covariance between n_2 and n_3 equals the extrinsic noise of either one

$$V_{23} = V_{11}\frac{H_{21}^2}{H_{22}^2}\frac{H_{22}/\tau_2}{H_{11}/\tau_1 + H_{22}/\tau_2}. \tag{S15}$$

As seen by comparing Eq. (S13) with the second term of Eq. (S14), and using the fact that components X_2 and X_3 have identical statistical properties, this is equivalent to

$$V_{23} = -V_{12}\frac{H_{21}}{H_{22}}. \tag{S16}$$

To prove it we have to solve the corresponding Eq. (S10) using the assumptions of equivalence and independence. However, since we are only interested in the relation between V_{12} and V_{23} we do not have to solve the entire system, only the equation defined by the 2nd rows and 3rd columns of the matrices in Eq. (S10). This directly proves Eq. (S16)

$$\frac{2}{\tau_2}\left(V_{12}H_{21} + V_{23}H_{22}\right) = 0. \tag{S17}$$

Because the underlying master equations are allowed to be nonlinear, it is tempting to think that this sanctifies the covariance strategy in experimental nonlinear systems. That is not the case. The approach above *ad hoc* assumes independence to establish first principles, but nonlinear systems rarely fulfil this criterion. Nonlinearities do not necessarily imply feedback loops or other forms of direct regulation. If the GFP molecules compete for proteases or if their transcripts compete for RNases, the degradation reactions become dependent and the method breaks down. Correlations may still say something useful, but the covariance will not coincide with the noise that comes from fluctuations in extrinsic factors. Fortunately, this will be evident in experiments, and was checked in the original study[6].

Autorepression, plasmid extinctions and transcription noise

Autorepression Adding negative feedback would typically decrease the average number of molecules and thereby increase intrinsic noise. This is an important issue in synthetic networks, but when analysing design principles of natural systems, it is often more convenient to compare them at the same average, i.e., assume that nature has provided compensatory mutations to restore the average[8]. The negative feedback loops in the experimental studies were artificially introduced or removed, but the averages are still not an issue. In the transcription study, the noise was unaffected by any changes in gene expression rates (including compensatory mutations), proving that these processes do not contribute substantial noise (see main text). The replication study did not look at one and the same system with and without negative feedback, but instead analysed a large number of natural plasmids and one unrelated synthetic plasmid where negative feedback was impaired. The average copy number of the synthetic plasmid was intermediate but relative fluctuations were still much higher.

There are also known exceptions from points (1)-(3) in section *Noise in the central dogma* of the main text, where autorepression increases rather than decreases the noise by introducing time-delays, noisy signalling or increasing the susceptibility to extrinsic changes. For instance, if the repression probability $K/(K+n_2)$ comes from binding and dissociating to DNA, the present stochastic treatment assumes rapid equilibration between bound and free DNA. If the equilibration instead is slow, the inherent switching stochasticity can introduce a large noise (see below). This problem is probably not significant in the experiments as expression was averaged over many plasmid copies[7], but the same autorepression of chromosomal gene expression may behave differently.

Plasmid extinctions Because self-replicators (e.g. plasmids) by definition are templates for their own synthesis, extinction eventually absorbs all probability mass. However, to approximate the distribution in cells where the process has not yet

gone to extinction, one can artificially introduce a small but nonzero escape rate, or obtain an exact equation by conditioning the process on not yet being extinct. For Markov processes, conditioning is done by simply re-routing the probability mass that enters the absorbing state to all other states proportionally. The approximations presented here only take into account how the system responds close to its average, and will thus approximate the distribution in plasmid-containing cells regardless, but extinctions and conditioning can be dealt with more explicitly[8,9].

Transcription noise In all the experimental systems, transcription was modified by changing the concentration of an inducer that targets a repressor (or sometimes an activator). This simultaneously affects many potential noise parameters: the fluctuations in the inducer concentration, the time-constant for de-activating the repressor, the repressor fluctuations, the time constant and average occupancy of the gene, the average transcription rate, and the average translation rate. Many of these changes could produce a decreasing protein noise with increasing inducer concentration, as observed in most experiments (see main text). The mRNA-protein model in section *Transcription and translation noise in B. subtilis* provides one example, but at least two other scenarios have been considered in the literature: noise from randomly turning on and off genes[10], and decreasing susceptibility to repressor noise[5].

The simplest gene model is the random telegraph process

$$on \underset{\mu}{\overset{\lambda}{\rightleftarrows}} off \ . \tag{S18}$$

For a single gene, the number of active genes n_1 then fluctuates between zero and one, and

$$\langle n_1 \rangle = \frac{\mu}{\mu+\lambda} \quad \text{and} \quad \frac{\sigma_1^2}{\langle n_1 \rangle^2} = \frac{1}{\langle n_1 \rangle} - 1 , \tag{S19}$$

where stationary n_1-fluctuations are Bernoulli distributed (as the outcome of flipping a biased coin). If proteins (X_2) are synthesised at a constant intensity when the gene is on and are eliminated through exponential decay, then

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} = \frac{1}{\langle n_2 \rangle} + \frac{\sigma_1^2}{\langle n_1 \rangle^2} \frac{(\mu+\lambda)^{-1}}{\tau_2 + (\mu+\lambda)^{-1}} \tag{S20}$$

similarly to Eq. (1) in the main text. By varying the repressor concentration, the experiments vary λ in Eq. (S18). Because the average protein level in this model is proportional to the average number of active genes, $\langle n_2 \rangle = A \langle n_1 \rangle$ where A is a constant, and because the experiments measure n_2, it is convenient to rewrite Eqs. (S19) and (S20) as

$$\frac{\sigma_1^2}{\langle n_1 \rangle^2} = \frac{A}{\langle n_2 \rangle} - 1 \qquad \text{and} \qquad \langle n_2 \rangle = A \langle n_1 \rangle = A \frac{\mu}{\mu+\lambda} \Rightarrow \frac{1}{\mu+\lambda} = \frac{\langle n_2 \rangle}{A\mu} \tag{S21}$$

Because the protein level additionally was so high that the first term in Eq. (S20) is negligible, protein noise follows an observational

$$\frac{\sigma_2^2}{\langle n_2 \rangle^2} \approx \left(\frac{A}{\langle n_2 \rangle} - 1 \right) \frac{\langle n_2 \rangle}{B + \langle n_2 \rangle} = \frac{A - \langle n_2 \rangle}{B + \langle n_2 \rangle} \tag{S22}$$

where B is a constant that depends on A, μ and τ_2, but not on λ that is varied in the experiments. Depending on parameters, this can provide an almost constant background noise that is independent of transcription, but can also be almost inversely proportional to transcription rate for certain parameters. It is thus possible that the 'mRNA' noise in the experimental interpretations in fact is a 'gene' noise, and the main support for the mRNA interpretation of the one-gene *B. subtilis* experiments[11] is that the gene model above only is consistent with the experimental results for particular parameter values, while the mRNA model needs no parameter fitting.

If the genes switch on and off independently of inducer, for example due to spontaneous conformational changes in the DNA, this could provide a constant source of noise that is independent of the repressor concentration. This may explain why the intrinsic noise in the two-gene *E. coli* study[6] (where the 'intrinsic' category would include gene noise) did not approach zero at high transcription rates, but rather fit $\sigma_2^2 / \langle n_2 \rangle^2 = c_1 + c_2 / \langle n_2 \rangle$ with $c_1 \approx 10^{-2}$ and $0 \le c_2 / \langle n_2 \rangle \le 0.3$, and thus $\sigma_2 / \langle n_2 \rangle \ge 10\%$. However, this observation could also be explained by the fact that the two GFPs were not perfectly identical, so that some extrinsic noise may be sorted as intrinsic.

If the protein fluctuations instead come from fluctuations in the repressor concentration, one must take into account that the susceptibility to repressor fluctuations changes with the average induction. For instance, assume that the corresponding rate equation is

$$\frac{\partial x_2}{\partial t} = k_2 \frac{K}{K + x_1} - d_2 x_2 \tag{S23}$$

where x_1 and x_2 are repressor and protein concentrations respectively. If all other things remain constant, the steady state concentration and susceptibility to repressor then follow

$$\frac{d_2}{k_2}x_2 = \frac{K}{K+x_1} \quad \text{and} \quad H_{21} = \frac{\partial}{\partial \ln x_1}\ln\left(\frac{d_2 x_2}{K/(K+x_1)}\right) = \frac{x_1}{K+x_1} = 1 - \frac{K}{K+x_1} = 1 - \frac{d_2}{k_2}x_2 = 1 - A\langle n_2\rangle \qquad \text{(S24)}$$

where constant A depends on d_2, k_2 and Ω. Using this in Eq. (1) thus again produces an observational decrease in $\sigma_2^2/\langle n_2\rangle^2$ with $\langle n_2\rangle$, though not an inverse proportionality throughout the induction range.

A non-monotonic $\sigma_2^2/\langle n_2\rangle^2$ could also come about in many ways. In the two-gene *E. coli* study[6] it was explained[5] by the lower susceptibility at higher induction (as in Eq. (S24)), that gradually becomes overshadowed by an increase in relative LacI fluctuations at lower LacI averages. But a lower average LacI would only have this effect if LacI fluctuations really come from having few LacI molecules per cell, not if LacI fluctuations come from fluctuations in mRNAs, plasmids or other factors that enslave LacI. An increase in relative LacI fluctuations may instead come from IPTG fluctuations, or, if dissociation of the LacI-IPTG complex is slow compared to its degradation and dilution rates, from the shorter effective LacI lifetime. If LacI transmits extrinsic noise from e.g. plasmids to GFP, a shorter lifetime reduces the total time-averaging in the cascade (Eq. (1)) and can reproduce the observed effect. However, it should be stressed that there is nothing plasmid-specific about this explanation and therefore no reason not to compare total noise in the two-gene *E. coli* study[6] with the total noise in the one-gene *S. cerevisiae* study[12], that instead attributed the non-monotonicity to transcriptional re-initiation.

Non-monotonicity also follows naturally from leaky gene expression. If a gene is expressed at a certain rate when a molecule X is bound, and at another rate when it is not bound, then gene expression will be constitutive at both low and high X. At intermediate X, however, the gene displays spontaneous fluctuations in activity, and additionally transmits fluctuations from X, typically producing non-monotonous susceptibilities. Both scenarios should be common in prokaryotes and eukaryotes alike, so non-monotonous noise curves can never be used to infer organismic differences, unless the experiments are carried out systematically for large numbers of genes.

Plasmid-dependent or independent fluctuations?

The plasmid-based experiments disagree on whether the noise comes from plasmids or other random cell processes. One possibility is that natural plasmids display e.g. 20% fluctuations, that the synthetic construct with impaired replication control displays e.g. 100%, and that the environment adds an intermediate 40%. If the normalised variances add, total noise would then be $\sqrt{0.2^2+0.4^2} = 44\%$ and $\sqrt{1^2+0.4^2} = 108\%$ respectively. GFP would then faithfully report fluctuations for the synthetic construct, but not for natural plasmids. The experimental controls give some further indications of where the noise comes from.

One indication of plasmid-independent noise is that all six natural plasmids examined in the replication study showed relative standard deviations in the range of 43% to 55%. Since the plasmids are unrelated, with different averages, control mechanisms and partitioning of copies at the end of the cell cycle, such similarities are surprising. However, it could be partly explained by the fact that low-copy plasmids compensate with tighter replication control and more precise partitioning at cell division, or that plasmids are subject to substantial extrinsic noise so that averages become irrelevant. Some of the similarity can also be explained by the fact that measurements were made over unsynchronised populations of cells, where dividing cells on average contain twice as many plasmid copies as new-born cells. If such systematic variation adds an extrinsic component, as naively expected for plasmids that regulate their concentration, then

$$\left(\frac{\sigma^2}{\langle n\rangle^2}\right)_{unsynch} = \left(\frac{\sigma^2}{\langle n\rangle^2}\right)_{synch} + C^2. \qquad \text{(S25)}$$

With C as the relative standard deviation in volume, the replication study indicated C = 21%, as expected theoretically[13] for unsynchronised cultures in idealised exponential growth, but the method used is expected to underestimate volume variation. If C =30%, the standard deviations for synchronised cells would range from 31% to 46% rather than from 43% to 55%. The transcription study[7] removed this source of variation by looking at the fluorescence per unit area of the cell profile, but two cloning vectors with a 20-fold difference in copy numbers still showed no significant differences in the noise levels.

If the plasmid measurement worked perfectly, cells with one plasmid copy would have half the fluorescence of cells with two plasmid copies. Since no cells have e.g. 1½ copies, these subpopulations should be clearly separable, at least for low-copy plasmids with averages of <5 per cell. But the observed fluorescence distributions were smooth and unimodal, without peaks at discrete copy numbers. This may indicate that the variation does not reflect gene dosage, at least not very accurately. However, the same method applied to chromosomes showed quite clear peaks. The difference could reflect time-averaging. For both plasmids and chromosomes, *gfp* was pulse-induced for 25% of a cell cycle, and the cells were fixed to let all the GFP molecules mature before measuring fluorescence. Thus the GFP level should depend on the recent

history of plasmid copy numbers – ideally the current number – rather than an average over the last couple of generations. However, the efficiency of time-averaging also depends on the mechanisms of replication control. Chromosomes replicate in synchrony, thus spending longer times at a certain copy number and giving the GFP level more time to adjust, while most bacterial plasmids do not replicate in synchrony. Pulse-inducing GFP synthesis for 25% of the cell cycle may thus be too long for plasmids, but shorter times create new problems: 1) fluctuations from having few mRNA transcripts start to contribute; 2) random association or dissociation of transcriptional activators or repressors significantly randomise the effective induction time; 3) other extrinsic fluctuations also become less time-averaged; and 4) the average GFP level was not proportional to induction time – the controls instead show a large constant term that will contribute more at short induction.

Supporting the interpretation that plasmid-encoded GFP reports plasmid copy numbers, some candidates for plasmid-independent noise can be ruled out. If the noise came from the probabilistic events in gene expression, high copy plasmids would display much smaller relative standard deviation since gene expression then would be averaged over more copies. That the relative standard deviation was unaffected by the average numbers of genes, mRNAs and proteins thus proves that the noise comes from other cell factors, like plasmids, RNA polymerase, ribosomes, proteases etc. The chromosomal studies of constitutive gene expression also showed low noise levels in spite of the lower number of genes. Finally, the extrinsic noise in the two-gene study was large when the transcriptional repressor was plasmid-coded. All these results indicate that much of the noise comes from plasmids, but more direct measurements are necessary.

In conclusion, GFP could both exaggerate and underestimate plasmid fluctuations, but there are certainly indications that plasmids fluctuate substantially. The plasmid replication study estimated that the GFP method was good enough to separate two-fold differences in copy numbers, which is consistent with the other studies.

1. N.G van Kampen *Stochastic Processes in Physics and Chemistry*. North-Holland (1992).
2. Elf, J. & Ehrenberg, M. Fast evaluations of fluctuations in biochemical networks with the linear noise approximation. *Genome. Res.* **13**, 2475-2484 (2003).
3. Keizer, K. *Statistical Thermodynamics of Nonequilibrium Processes*. Springer-Verlag, Berlin (1987).
4. Risken, H. *The Fokker-Planck Equation*. Springer-Verlag (1996).
5. Swain, P.S., Elowitz, M.B. & Siggia, E.D. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proc. natl. Acad. Sci. USA* **99**, 12795-12800 (2002).
6. Elowitz M.B., Levine A.J., Siggia E.D. & Swain P.S. Stochastic gene expression in a single cell. *Science* **297**, 183-186 (2002).
7. Becskei A. & Serrano L. Engineering stability in gene networks by autoregulation. *Nature* **405**, 590-593 (2000).
8. Paulsson, J. & Ehrenberg, M. Noise in a minimal regulatory network: plasmid copy number control. *Quart. Rev. Biophys*. **34**, 1-59 (2001).
9. Paulsson, J. & Ehrenberg, M. Random signal fluctuations can reduce random fluctuations of regulatory components of chemical regulatory networks. *Phys. Rev. Letters* **23**, 5447-5450 (2000).
10. Peccoud, J. & Ycart, B. Markovian modelling of gene-product synthesis. *Theor. pop. Biol.* **48**, 222-234 (1995).
11 Ozbudak E.M., Thattai, M., Kurtser, I., Grossman, A.D., & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nat*. *Genet*. **31**, 69-73 (2002).
12. Blake, W.J., Kærn M, Cantor, C.R. & Collins J.J. Noise in eukaryotic gene expression. *Nature* **422**, 633-7 (2003).
13. Cooper, S. & Helmstetter, C.E. Chromosome replication and the division cycle of Escherichia coli B/r. *J Mol. Biol.* **31**, 519-540 (1968).