Transcription stochasticity of complex gene regulation models

Anne Schwabe[1,2], Katja N. Rybakova[3], and Frank J. Bruggeman[1,2,3,4]

[1]Life Sciences, Centre for Mathematics and Computer Science (CWI), Science Park 123, 1098XG Amsterdam, The Netherlands
[2]Netherlands Institute for Systems Biology, Science Park 904, NL-1098 XH Amsterdam, The Netherlands
[3]Molecular Cell Physiology, Faculty of Earth and Life Sciences, Vrije Universiteit, De Boelelaan 1087, 1081HV Amsterdam, The Netherlands
[4]Swammerdam Institute for Life Sciences, Science Park 904, NL-1098 XH Amsterdam, The Netherlands

Supporting Material

A Molecular ratchet: reconstruction from the literature

Here we describe the reconstructed cyclic sequence of known nucleosome modifications in the transcription cycle found in yeast (figure S-1). It is well established that many yeast transcription factors such as Gcn4p and Gal4 interact with complexes SAGA and NuA4 that contain histone acetylases Gcn5 and Esa1, respectively (1, 2, 3). This leads to the acetylation of several sites in the H3 histone tail, namely H3K14, H3K9 and H3K18 by Gcn5, as well as acetylation of K5, K8 and K12 of the H4 tail by Esa1 (4). Acetylation of H3K14 has been shown to enhance binding and nucleosome-displacing activity of the remodeling complex SWI/SNF (5). Both the acetylation of histone tails and the removal of nucleosomes from the TATA-box of the promoter contribute to establishing the initiation-permissive state. It was demonstrated that the basal transcription factors Bdf1, which promotes binding of TFIID has high affinity for acetylated H4 (6), and that histone displacement accompanies transcription activation (7).

Once the promoter becomes accessible, the initiation complex containing polymerase is assembled; the elongation competent PolII is known to associate with the chromatin-modifying complex COMPASS responsible for the H3K4 mono-, di- and tri-methylation (8). These histone marks promote the deactivation of the promoter in several ways. H3K4 di-methylation has been shown to attract the Set3 complex (9) containing HOS2 and Stp1 enzymes capable of deacetylating the H3 and H4 lysine residues (10). Both di- and tri-methylated H3K4 facilitates the binding of the nucleosome-remodeling complex Isw1 (11), which has been implicated in limiting gene activation - one might speculate through reversing the remodeling that caused exposing the TATA-box. In order to complete the transcriptional cycle and return the promoter to the original state demethylation of H3K4 is required. An enzyme KDM1 has been demonstrated to have relevant activity(12), however, its specificity towards histone modifications has not been described so far.

B Waiting Time Distributions

As there is limited experimental information on the mechanisms of protein complex assembly, we investigated the effects of the assembly mechanism on the form of the transition waiting time distribution as well as of the reversibility of binding events

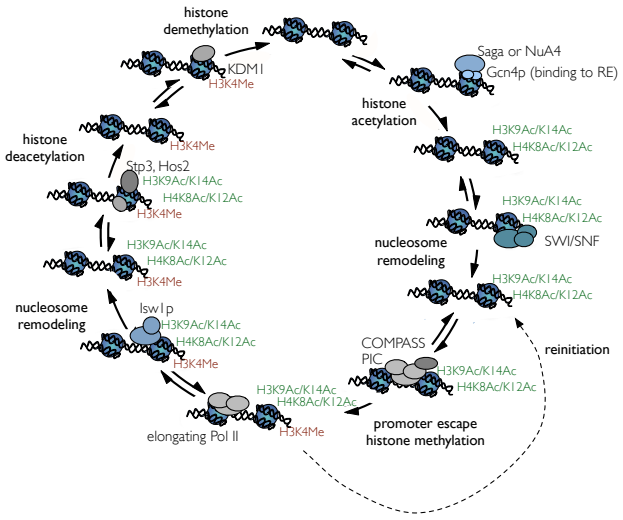


Figure S-1: **Core events of the transcription initiation cycle on a regulated yeast promoter.** The mechanism is a based on available experimental data for transcription factor and covalent histone modification mediated progression of transcription initiation in yeast.

and kinetic constants. We considered the assembly-time statistics of complexes of four proteins on chromatin; following either a completely sequential, a fully random, or one of out six different partially random mechanisms. The assembly schemes are shown in Figure S-2A. For the example of the partially random mechanism 2 (Fig S-2A) we demonstrate how the first passage time distribution for complex formation can be calculated in different ways.

Assuming that all reactions are irreversible, it is possible to derive this PDF as a convolution of PDFs for each binding step. The respective probability densities for each step are:

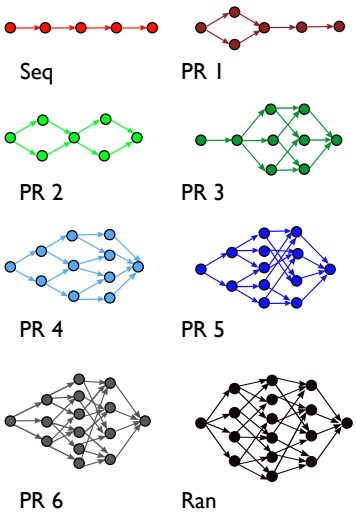
$$\begin{aligned} p_{step1}(t) &= (k_{1a}+k_{1b})e^{-(k_{1a}+k_{1b})t} \\ p_{step2}(t) &= \frac{k_{1a}}{k_{1a}+k_{1b}} \times k_{2a}e^{-k_{2a}t} + \frac{k_{1b}}{k_{1a}+k_{1b}} \times k_{2b}e^{-k_{2b}t} \\ p_{step3}(t) &= (k_{3a}+k_{3b})e^{-(k_{3a}+k_{3b})t} \\ p_{step4}(t) &= \frac{k_{3a}}{k_{3a}+k_{3b}} \times k_{4a}e^{-k_{4a}t} + \frac{k_{3b}}{k_{3a}+k_{3b}} \times k_{4b}e^{-k_{4b}t} \end{aligned} \tag{S-1}$$

The convolution of these PDFs equals the first passage time distribution of complex formation and can be expressed in form of its laplace transform:

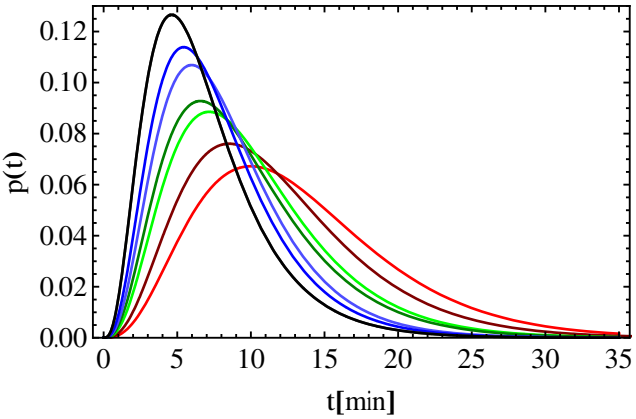
$$\mathcal{L}(p(t)) = f(s) = \mathcal{L}(p(step1)) \times \mathcal{L}(p(step2)) \times \mathcal{L}(p(step3)) \times \mathcal{L}(p(step4)) \tag{S-2}$$

The moments of the first passage time distribution are then given by

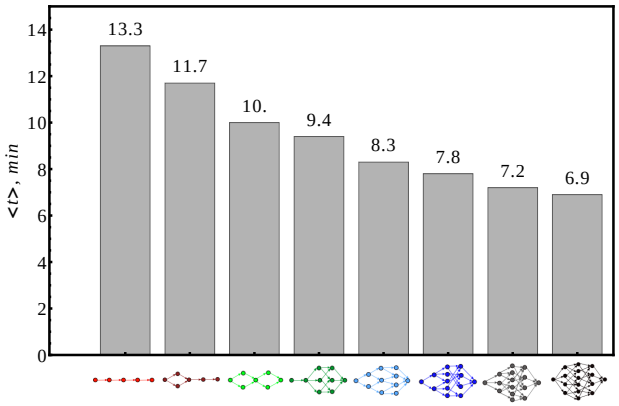
$$\langle t^a \rangle = (-1)^a \frac{d^a f(s)}{ds^a}|_{s=0}. \tag{S-3}$$



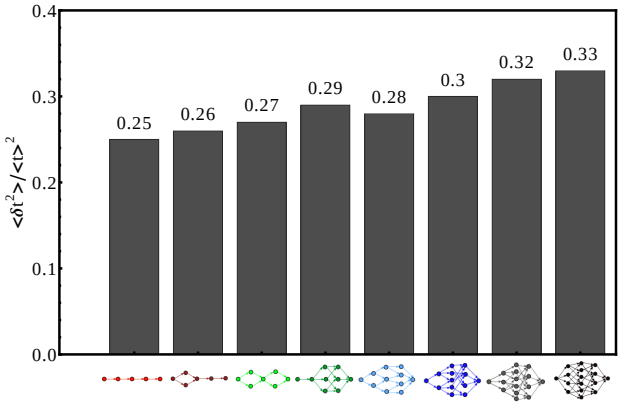
(a) Different assembly mechanisms



(b) Assembly time distributions



(c) Mean assembly times



(d) Noise in assembly time

Figure S-2: **Evaluation of assembly-time distributions for different four protein complex formation mechanisms** . a) Different assembly mechanisms: sequential (Seq), six partially random mechanisms (PR), and fully random (Ran). b) Assembly time distributions for the eight assembly mechanisms (modeled irreversible with equal effective rate constants $k = 0.3min^{-1}$: product of diffusion-limited association rate of $3.6nM^{-1}min^{-1}$ and 0.083nM protein concentration corresponding to 50 molecules in 1pL nuclear volume). c) Mean assembly times for the eight mechanisms. d) Noise in assembly time for the eight assembly mechanisms.

If all the rate constants are equal the full PDF can be determined:

$$p(t) = 4\exp^{-2kt} k \times (2 + kt + \exp^{kt} \times (kt - 2)) \tag{S-4}$$

Alternatively, the first passage time distribution can be described as a phase-type distribution. The continuos phase-type distribution for partially random mechanism 2 can be written as

$$\begin{aligned} p(t) &= \alpha e^{St} S_0 \\ S &= \begin{pmatrix} -k1a - k1b & k1a & k1b & 0 & 0 & 0 \\ 0 & -k2a & 0 & k2a & 0 & 0 \\ 0 & 0 & -k2b & k2b & 0 & 0 \\ 0 & 0 & 0 & -k3a - k3b & k3a & k3b \\ 0 & 0 & 0 & 0 & -k4a & 0 \\ 0 & 0 & 0 & 0 & 0 & -k4b \end{pmatrix} \\ \alpha &= (1,0,0,0,0,0,0,0)^T \\ S_0 &= S(1,1,1,1,1,1,1,1)^T, \end{aligned} \tag{S-5}$$

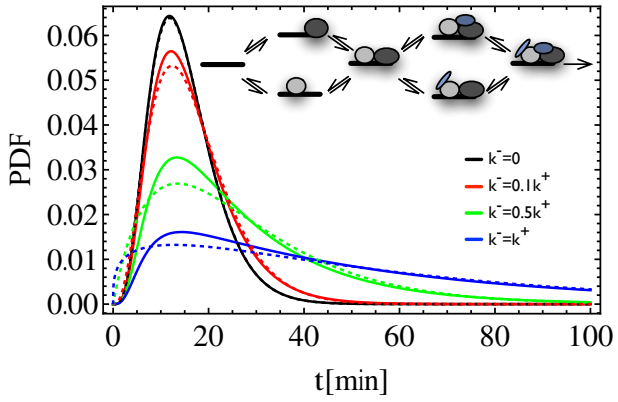


Figure S-3: **Reversibility causes long tails in first passage time distributions.** First passage time PDFs for the mechanism shown in the inset. Effective $k^+ = 0.3min^{-1}$, irreversible case (black) and k^+/k^- ratios 0.1 (red), 0.5 (green) and 1 (blue). Solid lines represent exact waiting time PDFs for the indicated mechanism, dashed lines correspond to gamma distributions with the same mean and variance.

with the first two moments given by:

$$\langle t \rangle = -\alpha S^{-1}\mathbf{1} \tag{S-6}$$

$$\langle t^2 \rangle = 2\alpha S^{-2}\mathbf{1} \tag{S-7}$$

The elements S_{ij} of the matrix S equal the rate constants for the reaction leading from state i to state j. This naturally allows to include reversible reactions. As shown in figures S-3 reversibility of reactions has a significant effect on the shape of the first passage time distribution if the apparent association rates are less than an order of magnitude larger than the dissociation rates.

However, even though reversibility can significantly change the first passage time distributions for individual ratchet transitions, the waiting time distributions for individual transitions as well as the convolution of several such distributions describing the distribution of the life time of an *on* or *off*-state can well be approximated by a gamma distribution that has the same mean and variance for the range of biologically relevant parameters. In simulations with uniformly random sampling of parameters for the assembly of protein complexes within ranges $10^{-4}s^{-1} \leq k^+ \leq 10^{0}s^{-1}$, $10^{-2}s^{-1} \leq k^- \leq 10^{-1}s^{-1}$ and $10^{-2}s^{-1} \leq k_{mod} \leq 10^{-1}s^{-1}$ the average KL-divergence index between the complete waiting time pdf and a gamma distribution with the same mean and variance was between $0.024 and 0.026$ for either single ratchet transitions or convolutions of 3 or 5 transitions with different mechanisms for complex assembly (sequential, or 3 different preferential random mechanisms). The low KL-indices show that for this parameter range resulting waiting time distributions for *on*- and *off*-states can well be approximated by gamma distributions, independent of the number of complexes that need to be formed on the DNA and independent of the details of the assembly mechanism.

C Burst-Size Distributions

For general distributions of $h(t)$, the waiting time distribution characterizing the time required for transcription initiation and promoter clearance, the moments of the burst-size distribution can be obtained in form of Laplace transforms. For

$T_{on} = t_{on}$:

$$\begin{aligned} p_b(B = b|T_{on} = t_{on}) &= p_b(B \geq b|T_{on} = t_{on}) - p_b(B \geq b+1|T_{on} = t_{on}) \\ &= \int_{t=0}^{t_{on}} (h(t)^b - h(t)^{(b+1)})dt \end{aligned} \tag{S-8}$$

The expectation value and variance of this distribution as a function of t_{on} are known in queuing theory as the renewal function and the variance function, respectively (e.g. (13)). For general distributions of $h(t)$, they can be expressed in terms of their Laplace transforms (14, 15):

$$\langle b(t)\rangle = \mathcal{L}^{-1}\left(\frac{\mathcal{L}(h(t))}{1-\mathcal{L}(h(t))}\right) \tag{S-9}$$

and

$$\langle \delta b(t)^2\rangle = \mathcal{L}^{-1}\left(\frac{\mathcal{L}(h(t))(1+\mathcal{L}(h(t)))}{(1-\mathcal{L}(h(t)))^2}\right). \tag{S-10}$$

The moments of the burst-size distribution can be obtained after integrating over all T_{on} equivalently to Eq. 1:

$$\begin{aligned} \langle b\rangle &= \int_{t=0}^{\infty} \langle b(t)\rangle g(t)dt \\ \langle b^2\rangle &= \int_{t=0}^{\infty} (\langle \delta b(t)\rangle + \langle b(t)\rangle^2) g(t)dt \end{aligned} \tag{S-11}$$

D mRNA noise for the ratchet model

D.1 Moments of steady-state mRNA distributions - Instantaneous burst model

First we consider a simplified model for a system with very pronounced bursts that can be approximated as producing a number of B transcripts instantaneously, where B is drawn from a general burst-size distribution. The time intervals between bursts are described by the distribution of the life times of the *off*-state and assumed to be independent of the burst-size (in queuing theory this description corresponds to a G/M/∞ queue with batch arrival). For exponentially distributed life times, the noise in mRNA levels for such a burst model can be described by the use of a stochastic hybrid system and Dynkin's formula (16) (which allows to solve for steady state moments as well as describing the time evolution of moments) or through the use of the chemical master equation in combination with moment equations (17) (exact solution for the steady-state mRNA noise and an approximation for systems with non-exponentially distributed life times of the *off* duration).

The method presented here consists of two steps: first, to use boundary conditions to determine the first k raw moments of the mRNA distribution at the onset and the end of a burst, and second to derive the moments of the full steady-state mRNA distribution from this. The evolution of the moments of the mRNA distribution during the *off*-state is described by degradation as the only process. Since each mRNA molecule has the same probability to be degraded within a certain time t (the life time distribution of mRNAs is exponential and therefore memoryless), the number of mRNAs after time t, given that no new burst has occurred yet, is distributed according to a binomial distribution, where the initial number of molecules that are present at the beginning of the *off*-state and the duration of the *off*-state are random variables themselves. Denote the moment generating function (mgf) of the copy number distribution at time t as $M(z,t)$ and $M(z,0) = M_i(z)$. Since each mRNA molecule has a survival probability of $e^{-k_d t}$, the mgf at time t can be expressed as a function of $M_i(z)$:

$$M(z,t) = M_i(Log[1 + e^{-k_d t}(e^z - 1)]) \tag{S-12}$$

and therefore at the end of the *off*-state:

$$M(z,t_{off}) = \int_{t=0}^{\infty} M_i(Log[1+e^{-k_d t}(e^z-1)])f(t)dt \tag{S-13}$$

Let the mgf of the burst-size be $M_{burst}(z)$, then the convolution of the probability distributions of molecules left from previous bursts and the probability distribution of a new burst are described by:

$$\begin{aligned} M(z,t_{burst}) &= M(z,0) = M_{burst}(z)M(z,t_{off}) \\ &= M_{burst}(z)\int_{t=0}^{\infty} M_i(Log[1+e^{-k_d t}(e^z-1)])f(t)dt \end{aligned} \tag{S-14}$$

This allows to determine the moments of the mRNA copy number at the time of a burst. The kth moment is given by

$$\langle n^k\rangle_i = \left.\frac{dM_i(z)}{dz}\right|_{z=0} \tag{S-15}$$

Taking the first k derivatives of $M_i(z)$ with respect to z yields a set of linear equations that can be solved for the first k moments of the mRNA copy number distribution at $t=0$:

$$\begin{aligned} \langle n\rangle_i &= \langle n_{burst}\rangle + \int_{t=0}^{\infty} e^{-k_d t}\langle n\rangle_i f(t)dt \\ \langle n^2\rangle_i &= \langle n_{burst}^2\rangle + \langle n_{burst}\rangle\langle n\rangle_i + \int_{t=0}^{\infty}(e^{-2k_d t}\langle n^2\rangle_i + e^{-k_d t}(1-e^{-k_d t})\langle n\rangle_i)f(t)dt \end{aligned} \tag{S-16}$$

Assuming a gamma distribution with shape parameter N_f and rate parameter k_f for the distribution of the life time of the *off*-state, $f(t)$, these moments are given by:

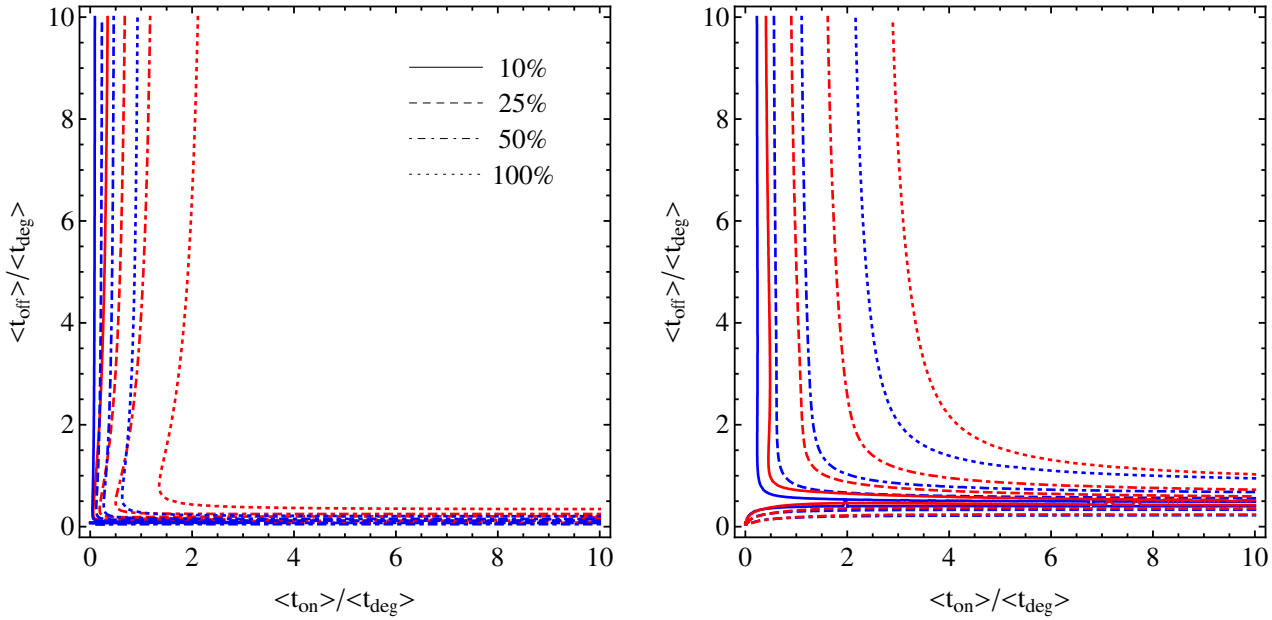
$$\begin{aligned} \langle n\rangle_i &= \frac{\langle n_{burst}\rangle}{1-\left(\frac{k_f}{k_f+k_d}\right)^{N_f}} \\ \langle n^2\rangle_i &= \frac{\langle n_{burst}^2\rangle + k^{N_f}\left((1+2\langle n_{burst}\rangle)(k_f+k_d)^{-N_f} - (k_f+2k_d)^{-N_f}\right)\langle n\rangle_i}{1-\left(\frac{k_f}{k_f+2k_d}\right)^{N_f}} \end{aligned} \tag{S-17}$$

With this, the moments at any time t during the *off*-state can be determined from Eq. S-12 as:

$$\begin{aligned} \langle n(t)\rangle &= \langle n\rangle_i e^{-k_d t} \\ \langle n^2(t)\rangle &= \langle n^2\rangle_i e^{-2k_d t} + e^{-k_d t}(1-e^{-k_d t})\langle n\rangle_i \end{aligned} \tag{S-18}$$

The moments of the full mRNA distribution can be obtained by averaging over the moments at times t weighted according to the probability that the next burst has not yet occurred at t, i.e. the survival probability $p_s(t) = 1-F(t)$ (with $F(t)$ as the cumulative distribution function of the distribution of *off*-state life times):

$$\begin{aligned} \langle n\rangle &= \frac{1}{\langle t_{off}\rangle}\int_{t=0}^{\infty}\langle n(t)\rangle(1-F(t))dt \\ \langle n^2\rangle &= \frac{1}{\langle t_{off}\rangle}\int_{t=0}^{\infty}\langle n^2(t)\rangle(1-F(t))dt \end{aligned} \tag{S-19}$$



(a) % deviation (exponential switch)

(b) % deviation (Erlang 5/5 switch)

Figure S-4: **For average life times of the *on*-state that are considerably shorter than the average degradation time, the instantaneous burst model is a good approximation of the gene switch modeled with gamma distributed life times.** The rate of transcription initiation k_{tr} was adjusted to keep the average steady-state mRNA level constant at $\langle n \rangle = 10$. Red lines indicate the use of the effective burst-size in the instantaneous burst model, whereas for the blue lines the regular burst-size was used. a) both *on* and *off*-state modeled with exponentially distributed life times. b) life times of *on* and *off*-state modeled with Erlang(5) distributions.

The solution to a full ratchet model with general duration distributions for *on*- and *off*-states in the following section allows to explore the range of applicability of the instantaneous burst model. Figure S-4, shows the percent deviation in the mRNA noise for an instantaneous burst model as compared to the full model as a function of the average durations for *on*- and *off*-states. The approximation is valid when the average duration of the *on*-state is shorter than the average life time of mRNA while the ratio of the average durations of *on* and *off*-states has a relatively small effect on the approximations validity. The approximation is improved when the effective burst-size instead of the regular burst-size is used. The deviations are larger for an exponential switch than for peaked distributions for the durations of both states.

D.2 Moments of steady-state mRNA distributions - general waiting time distributions for $f(t)$ and $g(t)$

The simplified instantaneous burst model can be extended to one that explicitly takes into account the duration(distribution) of the *on*-state as well: transcription initiation and mRNA degradation are described as first-order processes with rate constants k_h and k_d, and the distributions for the life times of *on* and *off*-states are described by general waiting time distributions. The reason for modeling transcription initiation with an exponential distribution is that the exact distribution for transcription initiation has a minor effect on the burst-size distribution and is therefore also expected to have little

effect on the steady state mRNA distribution. However, the approach described here only requires that the moments of birth-death process occurring during the *on*-state as a function of time are known; for general waiting time distributions of transcription initiation and promoter clearance, moments of this distribution can be calculated recursively (18). Depending on the time scales, the waiting time distribution for degradation of mRNA can have large effects, that can be difficult to predict intuitively. However, many measurements of mRNA turnover in mammalian cells can well be fit with an exponential decay model (19).

Let time $t = 0$ be the beginning of an *off*-state and $M(z,t)$ be the mgf of the number of mRNA molecules at time t ($0 \leq t \leq t_{on} + t_{off}$). Completion of one cycle (*off*- and *on*-state) defines the boundary condition

$$M(z,0) = M(z, t_{on} + t_{off}) = \int_{s=0}^{\infty} \int_{t=s}^{\infty} M(z,t) g(t-s) f(s) ds dt. \tag{S-20}$$

With the assumption of exponential life times of mRNAs the probability that one molecule is still present after time Δt is a Bernoulli random variable with mgf $h(z,t)$:

$$h(z, \Delta t) = 1 + e^{-k_d \Delta t}(e^z - 1). \tag{S-21}$$

Starting with the generating function $M(z,0) = M_i(z)$ and considering only degradation of mRNAs, at time t the compound mgf equals

$$M(z,t) = M_i(Log[h(z,t)]). \tag{S-22}$$

Therefore, at the end of an *off*-state the moment generating function is given by:

$$M(z, t_{off}) = \int_{t=0}^{\infty} M_i(Log[h(z,t)]) f(t) dt. \tag{S-23}$$

During an *on*-period degradation of the mRNA molecules that have been present at the beginning of the burst continues as described through the function $h(z,t)$ with $M_{deg}(z, t_{off} + u) = M_i(Log[h(z, u + t_{off})])$. In addition mRNAs are also created and degraded according to a birth-death-process with exponential waiting times. This can be described as an *effective* burst-size, B_{eff}, since this term only takes into account those mRNA molecules synthesized that survive at least until time u. The transient distribution for this birth-death process as described in (20) is given by:

$$p_{birth/death}(B_{eff} = b_{eff}, u) = \frac{1}{b_{eff}!} \left(\frac{k_h}{k_d} \left(1 - e^{-k_d u}\right) \right)^{b_{eff}} \left(e^{\frac{k_h}{k_d}\left(e^{(-k_d u)} - 1\right)} \right) \tag{S-24}$$

It is a Poisson distribution with average $\lambda_u = k_h/k_d(1 - e^{-k_d u})$ and mgf $M_{burst}(z, t_{off} + u) = e^{\lambda_u (e^z - 1)}$. During an *on*-state the probability distribution of the mRNA copy number distribution is given by the convolution of the distributions of molecules that are still present from previous bursts and the effective burst-size distribution. The generating function equals the product of the two individual generating functions:

$$\begin{aligned} M(z, t_{off} + u) &= M_{deg}(z, t_{off} + u) M_{burst}(z, t_{off} + u) \\ &= \int_{s=0}^{\infty} M_i(Log[1 + e^{-k_d(s+u)}(e^z - 1)) e^{k_h/k_d(1 - e^{-k_d u}(e^z - 1))} f(s) ds \end{aligned} \tag{S-25}$$

With Eq. S-20:

$$\begin{aligned} M(z, t_{off} + t_{on}) &= M(z,0) = M_i(z) \\ &= \int_{s=0}^{\infty} \int_{u=0}^{\infty} M_i(Log[1 + e^{-k_d(s+u)}(e^z - 1)]) e^{k_h/k_d(1 - e^{-k_d u}(e^z - 1))} f(s) g(u) ds du \\ &= \int_{s=0}^{\infty} \int_{t=s}^{\infty} M_i(Log[1 + e^{-k_d(t)}(e^z - 1)]) e^{k_h/k_d(1 - e^{-k_d(t-s)}(e^z - 1))} f(s) g(t-s) ds dt \end{aligned} \tag{S-26}$$

This integral equation is difficult to solve for general waiting time distributions, g and f, however the moments at $t = 0$ can be solved by differentiation of Eq. S-26. Taking the first k derivatives of $M(z, 0)$ with respect to z yields a set of linear equations that can be solved for the first k moments of the mRNA copy number distribution at $t = 0$:

$$\begin{aligned} \langle n \rangle_i &= \frac{\int_{u=0}^{\infty} \lambda_u g(u) du}{1 - \int_{s=0}^{\infty} \int_{u=0}^{\infty} e^{-k_d(s+u)} g(u) f(s) du ds} \\ \langle n^2 \rangle_i &= \frac{\int_{u=0}^{\infty} \int_{s=0}^{\infty} (\lambda_u(1+\lambda_u) + \langle n \rangle_i (e^{-k_d(u+s)} - e^{-2k_d(u+s)} + 2e^{-k_d(u+s)}\lambda_u)) g(u) f(s) du ds}{1 - \int_{s=0}^{\infty} \int_{u=0}^{\infty} e^{-2k_d(s+u)} g(u) f(s) du ds} \end{aligned} \tag{S-27}$$

The moments at any time t in the transcriptional cycle are then given by the derivatives of Eqs.s S-22 and S-25. To obtain the moments of the total mRNA copy number distribution one needs to average over all times according to their probabilities:

$$\langle n^k \rangle = \frac{1}{\langle t_{on} \rangle + \langle t_{off} \rangle} \left(\int_{s=0}^{\infty} \langle n^k \rangle_s (1 - F(s)) ds + \int_{u=0}^{\infty} \langle n^k \rangle_u (1 - G(u)) du \right), \tag{S-28}$$

with $\langle n^k \rangle_s$ as the kth moment of the mRNA copy number distribution at time s during the *off*-state (given by derivatives of Eq. S-22) and $\langle n^k \rangle_u$ as the kth moment at time $t = t_{off} + u$ during the *on*-state (derivatives of Eq. S-25) and F and G as the cumulative distribution functions of the duration distributions of *off* and *on*-states, respectively.

Figure S-5 shows a contour plot of the noise in the mRNA copy number for models where the waiting time distributions were modeled as either exponential or as Erlang distributions (four reactions in sequence). When the durations of the *on* and *off* state were varied, the average time between mRNA synthesis events was adjusted to keep the average transcript level constant at 10 molecules/cell. Figure S-5 indicates that multistep mechanisms in gene switching as well as in the initiation mechanism does reduce noise, but they have their largest effects in different parameter regimes. While multistep switching between *on* and *off* states leads to significant noise reduction in a burst regime (i.e. long *off* states interspersed with short periods of transcriptional activity), it's effects in a regime where the gene is *on* most of the time is less pronounced. In contrast, the multistep initiation mechanism has it's largest effect for a constitutively expressed gene where it reduces the noise in the mRNA distribution below the Poisson limit. In a bursting regime, the multistep initiation mechanism has almost no effect at all on the mRNA distribution as was already suggested by the analysis of noise in burst sizes. In a regime where the gene switches rarely to the *off* state and spends most of the time in the *on* state, the absolute magnitude of the noise reduction caused by either a multistep mechanism in the initiation or in the switching becomes comparable. Figure S-5C shows the distributions in this regime. Although the model with multistep switching and one step initiation mechanism has the exact same noise in mRNA copy numbers (black line) as a model with multistep initiation and single step switching (blue dashed line), the two distributions differ: the multistep switching mechanism leads to an almost zero probability of having no mRNA molecules in a cell but the distribution has a somewhat broader tail to the right in comparison with the distribution for the multistep initiation/single step switching model. This analysis suggests that depending on whether a gene is highly active, bursty, or constitutively active different aspects of the design of transcription determine the noise in transcription.

D.3 Steady State mRNA Distribution of a "Deterministic Switch"

As a limit to very precise life-time distributions for both states, these durations can be considered deterministic. The noise in the steady-state mRNA distribution then derives from the birth-death process that occurs during the *on*-state and the ongoing degradation during the *off*-state. Modeling synthesis and degradation of mRNA again with exponential waiting times again, the complete steady-state mRNA distribution of this 'deterministic switch' can be solved analytically. With

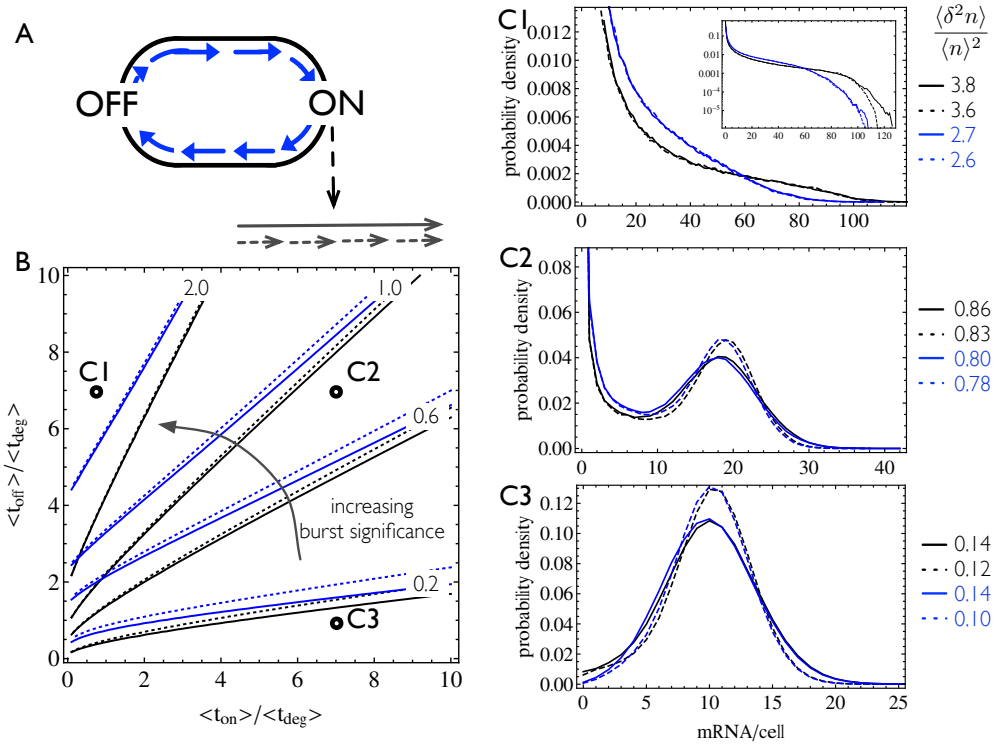


Figure S-5: **Multistep mechanisms in switching between *on* and *off* states as well as in the initiation mechanism itself can reduce noise in mRNA copy numbers** A. Cartoon depicting the different models used: blue lines indicate a four step mechanism for switching between *on* and *off* states, black lines a single step. Dashed lines are used for models with a four step mechanism for transcription initiation, while continuous lines are used for models with one step initiation. B. Contour plot of the noise in mRNA copy number distributions for the four models depicted in A for varying *on* and *off* state durations. The average time for transcription initiation was adjusted to keep the mean transcript level fixed at 10 molecules. Contours shown are for noise levels of 0.2, 0.6, 1, and 2. C. Copy number distributions for the three points indicated in B obtained through simulations.

deterministic times for both states Eq. S-26 simplifies to:

$$\begin{aligned} M(z, t_{off} + t_{on}) &= M(z, 0) \\ &= M(Log[1 + e^{-k_d(t)}(e^z - 1)])e^{k_h/k_d(1-e^{-k_d(t-s)})(e^z-1)} \end{aligned} \tag{S-29}$$

rewriting this with probability generating functions, $P(z,t) = M(Log[z], t)$ yields:

$$P(z, 0) = P_i(z) = P_i(1 + e^{-k_d t}(z - 1))e^{k_h/k_d(1-e^{-k_d(t-t_{off})})(z-1)}, \tag{S-30}$$

For $t = t_{on} + t_{off}$ this equation can be solved by iteration:

$$\begin{aligned} P(z, t_{on} + t_{off}) &= P(z, 0) \\ &= P_i(1 + e^{-k_d(t_{on}+t_{off})}(z - 1)])e^{k_h/k_d(1-e^{-k_d(t_{on})})(z-1)} \\ &= \underbrace{P_i(1, 0)}_{=1} \prod_{n=0}^{\infty} e^{k_h/k_d(1-e^{-k_d(t_{on})})e^{k_d t_{off}}(1+e^{-nk_d(t_{on}+t_{off})}(z-1))} \\ &= e^{\frac{e^{k_d t_{off}}\left(-1+e^{k_d t_{on}}\right)k_h}{\left(-1+e^{k_d(t_{off}+t_{on})}\right)k_d}(-1+z)}, \end{aligned} \tag{S-31}$$

which is the pgf of a Poisson distribution. During the *off*-state ($t < t_{off}$), $P(z,t)$ evolves according to:

$$P(z, t) = e^{\frac{e^{k_d(t_{off}-t)}\left(-1+e^{k_d t_{on}}\right)k_h}{\left(-1+e^{k_d(t_{off}+t_{on})}\right)k_d}(-1+z)} = e^{\lambda_{off}(t)(z-1)}, \tag{S-32}$$

while during the *on*-state

$$P(z, t) = e^{\frac{e^{k_d(t_{off}-t)}\left(-1+e^{k_d t_{on}}\right)k_h}{\left(-1+e^{k_d(t_{off}+t_{on})}\right)k_d}(-1+z)} \times e^{k_h/k_d(1-e^{-k_d(t-t_{off})})(z-1)} = e^{\lambda_{on}(t)(z-1)}, \tag{S-33}$$

Therefore, the distribution at any time t is given by a Poisson distribution with rate parameter $\lambda_i(t)$ (equal to the average mRNA copy number at time t), changing throughout the transcription cycle. The complete mRNA copy number distribution is given as the time average over the Poisson distributions with $\lambda(t)$:

$$P_{mRNA}(n) = \frac{1}{t_{on} + t_{off}} \left(\int_{s=0}^{t_{off}} \frac{\lambda_{off}(s)^n e^{-\lambda_{off}(s)}}{n!} ds + \int_{u=0}^{t_{on}} \frac{\lambda_{on}(u)^n e^{-\lambda_{on}(u)}}{n!} du \right) \tag{S-34}$$

D.4 Comparison of the different switch models

Since the exponential and deterministic switch models are opposing limits to models with Erlang (or gamma) distributions for the durations of the *on* and *off*-states, the steady-state mRNA distribution of any Erlang switch model can be expected to be an intermediate between the two distributions for the exponential and deterministic switch models, that have the same average life time for each state as well as the same rate constants for transcription initiation and mRNA degradation. For three different Erlang switch models, each with a total of $N_g + N_f = 10$ steps and eight different combinations of t_{on} and t_{off} that exemplify the different possible shapes of mRNA distributions (Fig. S-6), the steady-state mRNA distributions were obtained from simulations and compared to the distributions for the exponential and deterministic switch in terms of the Kullback-Leibler divergence index (Table S-1).

With at least five steps for each transition the resulting mRNA distribution is more similar to the distribution generated by a system with deterministic switching times than that of a system with exponential switching times. If one of the transitions occurs in only one step, the resulting distribution is much better described by an exponential than by a deterministic system. Similar results were obtained with other measures for the similarity of distributions then the KL-divergence index (results not shown).

D.5 Detailed analysis of Figure S-6B

As can be seen from figure S-6B, the differences between the two mRNA distributions are smallest if either t_{on} is much larger than t_{off} (the system is almost continuously in the *on*-state) or if t_{on} and t_{off} are of comparable size and both larger than the average time for degradation. The former case can be understood intuitively, this parameter combination leads to a distribution that is very similar to a Poisson distribution; the infrequent and short excursions to the *off*-state have only a small effect on the distribution and the effect of whether these switches to and from the *off*-state occur in more or less regular time intervals is also rather small.

An increase in the life time of *off*-state leads to a bimodal distribution. For the switch with exponential waiting times, the first peak of the bimodal distribution always lies at zero mRNA molecules. This is not the case for the deterministic switch in the region of large t_{on} and intermediate t_{off}. Accordingly, the Kullback-Leibler divergence index is relatively high in this region. A further increase of the life times of *off*-state also causes the deterministic switch to have its first peak at zero mRNA molecules, which is why in this regime the distributions for deterministic and exponential switch become more similar again (Fig S-6B). The regularity of the timing of switching between the states has a minor effect on the mRNA distribution in this regime. If the average time of the *on*-state is smaller than that for degradation, the distributions of the exponential and the deterministic switch are dissimilar for a large range of values of $\langle t_{off} \rangle$. For $\langle t_{off} \rangle \leq 1/k_d$ the mRNA distribution is a single-peaked distribution. In this regime, the noise in the waiting time distributions for the *on* and *off*-state have a large effect on the steady state mRNA distribution; the distribution for the exponential switch is much broader than that for the deterministic switch (Figure S-6B; see Fig. S-7 for example distributions). Greater durations of the *off*-state while keeping $\langle t_{on} \rangle$ smaller than $1/k_d$ makes the two distributions even more dissimilar. While the distribution for the exponential switch is now in a regime that can be described with a power law (21), the distribution of the deterministic switch still has its maximum at $n > 0$. Increasing t_{off} even more, i.e., increasing the burst-size, leads to a regime where also the deterministic switch has a distribution with its maximum at zero mRNA molecules. Nevertheless, the two distributions are still clearly distinguishable over a wide range of values for $\langle t_{off} \rangle$.

E Comparison to Experimental Data

The mRNA distributions measured by Zenklusen et al. (22) were fit to different Erlang switch models ($N_g = 1, N_f = 1$: red, $N_g = 1, N_f = 10$: blue, $N_g = 5, N_f = 5$: green, $N_g = 10, N_f = 1$: magenta, $N_g = 10, N_f = 10$: orange) by calculating setting $k_d = 1$ and choosing k_h, k_f, and k_g to give rise to a steady-state mRNA distribution with the first three moments equal to the experimentally observed distribution. The full mRNA distributions at steady-state were then obtained for these models through simulations. The parameters used for these simulations can be retrieved from Tables S-2 and S-3 ($k_f = N_f(1 + \langle t_{on} \rangle)/\langle t_{off} \rangle(\langle t_{on} \rangle + \langle t_{off} \rangle), k_g = N_g \langle t_{on} \rangle, k_h = \langle b \rangle / \langle t_{on} \rangle$). Even though the steady-state distributions are remarkably similar for all models, the underlying mechanisms, burst-sizes and time traces differ significantly.

Figure S-8 demonstrates, that for genes that display significant bursting, the waiting time distributions of the underlying mechanism could in principle be determined from data on the times of synthesis events (as would be observed during a mRNA-counting experiment). The simulated synthesis times from Figure 6c ($N_g = 1$, $N_f = 10$) were used to determine the average burst-size ($\langle b \rangle$) and duration (τ_b) from the sequence size function (23). The sequence size function can be used to determine the mean burst-size from a time series. All synthesis events separated by a period longer than the mean burst

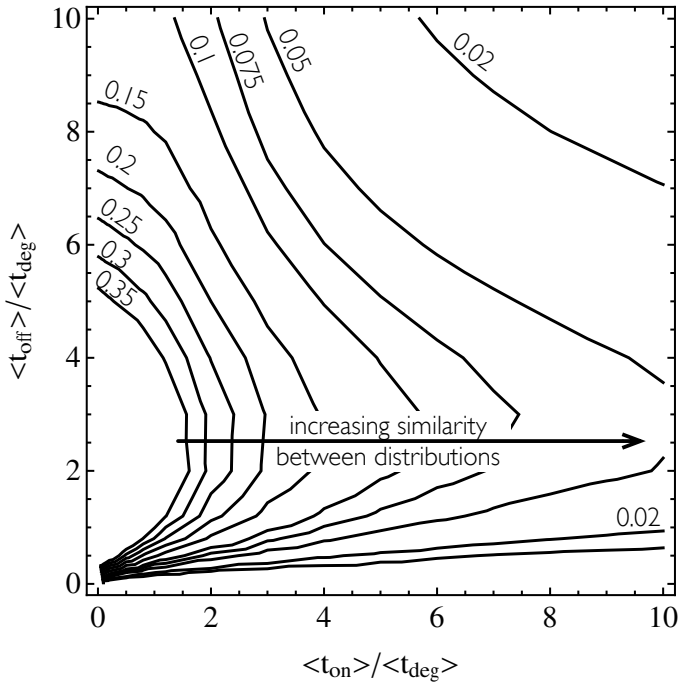


Figure S-6: **Timescale separation in the model with genetic switch, transcription and degradation largely determines the shape of the steady-state mRNA distribution.** The Kullback-Leibler divergence index (in bits) between the steady state mRNA distributions of the deterministic and the exponential switch shows that the effect of the precision of the duration of the states has the largest effect on the mRNA distribution in an intermediate regime of $\langle t_{on}\rangle/\langle t_{deg}\rangle$ vs. $\langle t_{off}\rangle/\langle t_{deg}\rangle$. In the exponential switching model: $k_f = \langle t_{off}\rangle^{-1}$ and $k_g = \langle t_{on}\rangle^{-1}$. Contours are shown for 0.01, 0.02, 0.05, 0.075, 0.1, 0.15, 0.2, 0.25, 0.3 and 0.35. The mean mRNA level, $\langle n\rangle$, was kept at 10 by adjusting the rate for transcription initiation, k_h. The degradation rate constant only changes the scaling of the plot and was therefore set to 1. Usage of other indices than the Kullback-Leibler index gave qualitatively similar results.

duration, τ_b, were scored as different *on*-phases. As shown in Figure S-8a, this classification of *on* and *off*-states based on the burst duration is correct more than 95% of the time. Estimates of the distribution functions g and f (describing the durations of *on* and *off*-states, respectively) are very similar to the actual distributions used to simulate the time-traces (shown in Figure S-8c and d).

The advantage of using the sequence size function (23) to identify the *on*-phases from the time-trace data is that no assumptions are made about the shape of g, f and h. Once the *on* and *off*-phases are identified, estimates for these distributions can then directly be obtained from the time-trace data. Since *on*-phases which do not lead to initiation events can not be detected and since the beginning and end of an *on*-phase are assigned to the time of the first and the last initiation event of the phase, this methods tends to a slight underestimate of the duration of *on*-phases and the mean burst-size and an overestimate of the duration of the *off*-phase. However, these effects are negligible if *on*-phases lead to many initiation events.

Table S-1: KL-divergence indices for the steady state mRNA distributions shown in figure S-7 and either a system with deterministic switching times (det) or one with exponential switching times (exp). The total number of steps for both transitions equals ten for all three systems.

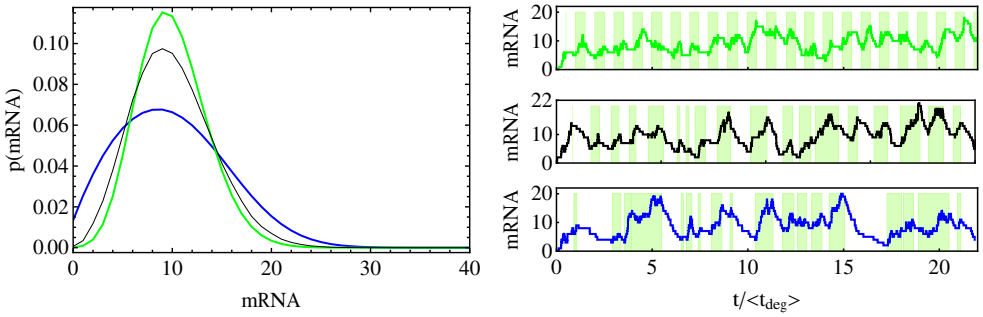
$\langle t_{on}\rangle/\langle t_{off}\rangle$	0.5/0.5	0.5/2.5	0.5/5.0	0.5/8.0	4.0/0.5	4.0/2.5	4.0/5.0	4.0/8.0
$D_{KL}(N_g=1,N_f=9,exp)$	0.04	0.11	0.07	0.04	0.01	0.05	0.04	0.02
$D_{KL}(N_g=1,N_f=9,det)$	0.20	0.67	0.56	0.49	0.01	0.04	0.03	0.01
$D_{KL}(N_g=9,N_f=1,exp)$	0.03	0.06	0.04	0.04	0.00	0.01	0.01	0.01
$D_{KL}(N_g=9,N_f=1,det)$	0.24	0.72	0.33	0.15	0.03	0.16	0.07	0.04
$D_{KL}(N_g=5,N_f=5,exp)$	0.15	0.23	0.14	0.07	0.01	0.07	0.05	0.03
$D_{KL}(N_g=5,N_f=5,det)$	0.04	0.19	0.11	0.07	0.00	0.02	0.01	0.00

Table S-2: Comparison of models with different numbers of steps in the transition between *on* and *off* states (N_g,N_f) for PDR5 and MDN1.

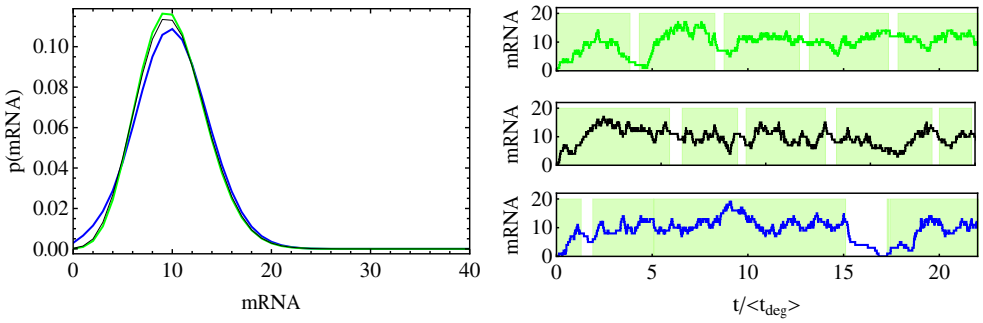
		PDR5			MDN1		
N_g	N_f	$\langle b\rangle$	$\frac{\langle t_{on}\rangle}{\langle t_{off}\rangle}$	$\frac{\langle t_{on}\rangle+\langle t_{off}\rangle}{\langle t_{deg}\rangle}$	$\langle b\rangle$	$\frac{\langle t_{on}\rangle}{\langle t_{off}\rangle}$	$\frac{\langle t_{on}\rangle+\langle t_{off}\rangle}{\langle t_{deg}\rangle}$
1	1	4.8	0.0215	0.36	1.6	2.1	0.28
1	10	12.4	0.2364	0.93	5.2	3.0	0.86
5	5	16.2	0.0007	1.20	6.4	2.2	1.05
10	1	8.6	0.0009	0.64	1.8	1.4	0.32
10	10	21.4	0.0005	1.60	9.0	2.3	1.48

Table S-3: Comparison of models with different numbers of steps in the transition between *on* and *off* states (N_g,N_f) for POL1, DOA1 and KAP104.

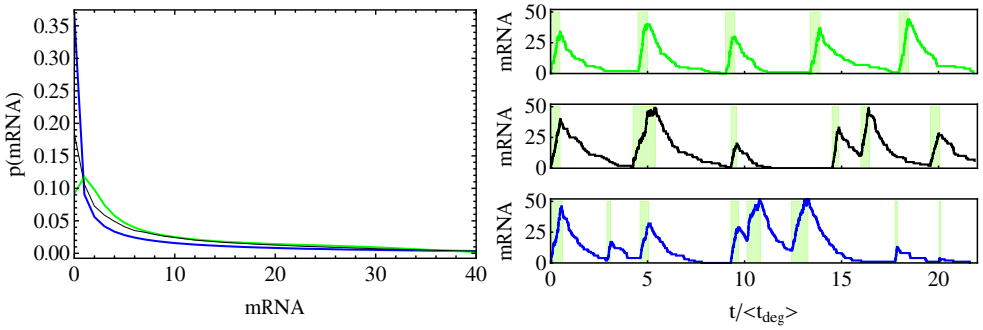
		POL1			DOA1			KAP104		
N_g	N_f	$\langle b\rangle$	$\frac{\langle t_{on}\rangle}{\langle t_{off}\rangle}$	$\frac{\langle t_{on}\rangle+\langle t_{off}\rangle}{\langle t_{deg}\rangle}$	$\langle b\rangle$	$\frac{\langle t_{on}\rangle}{\langle t_{off}\rangle}$	$\frac{\langle t_{on}\rangle+\langle t_{off}\rangle}{\langle t_{deg}\rangle}$	$\langle b\rangle$	$\frac{\langle t_{on}\rangle}{\langle t_{off}\rangle}$	$\frac{\langle t_{on}\rangle+\langle t_{off}\rangle}{\langle t_{deg}\rangle}$
1	1	5.7	0.20	1.81	0.14	0.010	0.0620	13.6	6.7	2.84
1	10	10.1	0.29	3.23	0.24	0.0010	1.0	35.2	8.3	7.32
5	5	11.4	0.14	3.61	0.62	0.0004	2.6	34.1	7.5	7.08
10	1	6.4	0.05	2.05	0.25	0.0012	1.1	13.4	5.6	2.79
10	10	11.9	0.09	3.77	0.99	0.00002	4.1	40.3	7.9	8.38



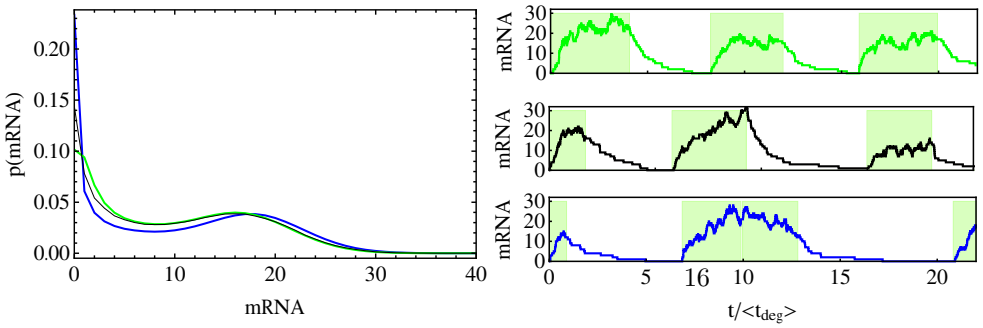
(a) $\langle t_{on}\rangle/\langle t_{deg}\rangle = 0.5$, $\langle t_{off}\rangle/\langle t_{deg}\rangle = 0.5$



(b) $\langle t_{on}\rangle/\langle t_{deg}\rangle = 4$, $\langle t_{off}\rangle/\langle t_{deg}\rangle = 0.5$

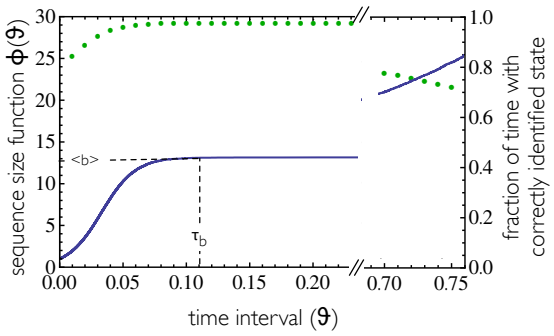


(c) $\langle t_{on}\rangle/\langle t_{deg}\rangle = 0.5$, $\langle t_{off}\rangle/\langle t_{deg}\rangle = 4$

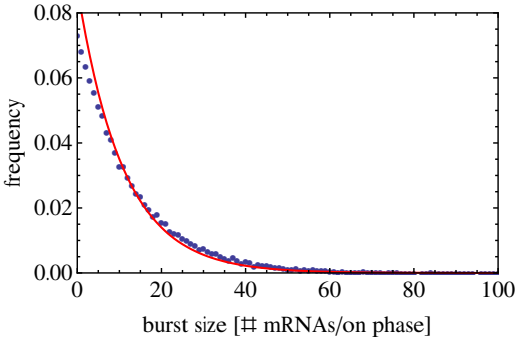


(d) $\langle t_{on}\rangle/\langle t_{deg}\rangle = 4$, $\langle t_{off}\rangle/\langle t_{deg}\rangle = 4$

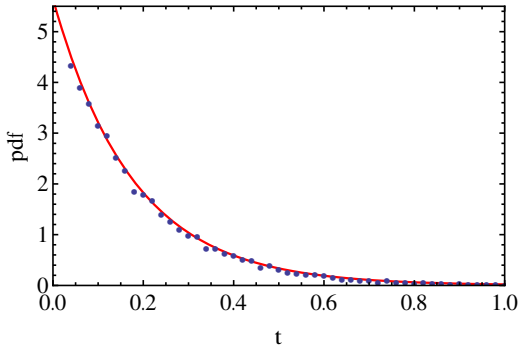
Figure S-7: Steady state mRNA distribution for different parameter regimes and different waiting time distributions for switching: BLUE: exponential waiting times, GREEN: deterministic waiting times, BLACK: N_g : 5, N_f : 5. $\langle n\rangle = 10$, $k_d = 1$, $\langle t_{on}\rangle$ and $\langle t_{off}\rangle$ as indicated for each plot, k_f and k_g were adjusted accordingly.



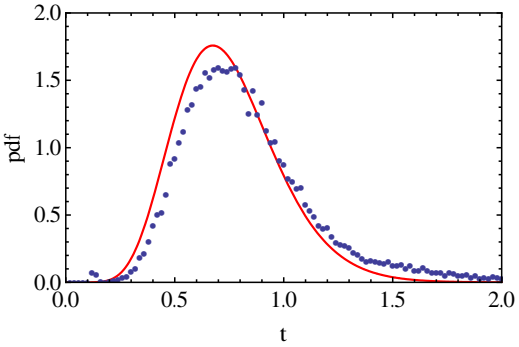
(a)



(b)



(c) *on* time distribution



(d) *off* time distribution

Figure S-8: **The life-time statistics of the switch can be obtained from time-trace data of mRNA.** A) The sequence size function of the stochastic time trace (blue) and the fraction of time that is correctly classified as belonging to an *on* or *off*-phase where the classification is based on the time scale (υ; dimensionless time as unit) (green). B-D: Distributions obtained from a classification into *on* and *off*-states based on τ_b (blue) and the exact distributions (red). B) burst-size distribution. C) First-passage-time pdf of the *on*-state duration, $g(t)$. D) First-passage-time pdf of the *off*-state duration, $f(t)$.

References

[1] Drysdale, C. M., B. M. Jackson, R. McVeigh, E. R. Klebanow, Y. Bai, T. Kokubo, M. Swanson, Y. Nakatani, P. A. Weil, A. G. Hinnebusch, 1998. The Gcn4p activation domain interacts specifically in vitro with RNA polymerase II holoenzyme, TFIID, and the Adap-Gcn5p coactivator complex. *Mol. Cell. Biol.*, 18:1711–1724.

[2] Larschan, E., F. Winston, 2005. The Saccharomyces cerevisiae Srb8-Srb11 complex functions with the SAGA complex during Gal4-activated transcription. *Mol. Cell. Biol.*, 25:114–123.

[3] Herbig, E., L. Warfield, L. Fish, J. Fishburn, B. A. Knutson, B. Moorefield, D. Pacheco, S. Hahn, 2010. Mechanism of Mediator recruitment by tandem Gcn4 activation domains and three Gal11 activator-binding domains. *Mol. Cell. Biol.*, 30:2376–2390.

[4] Mitarai, N., K. Sneppen, S. Pedersen, 2008. Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J. Mol. Biol.*, 382:236–245.

[5] Chandy, M., J. L. Gutierrez, P. Prochasson, J. L. Workman, 2006. SWI/SNF displaces SAGA-acetylated nucleosomes. *Eukaryotic Cell*, 5:1738–1747.

[6] Durant, M., B. F. Pugh, 2006. Genome-wide relationships between TAF1 and histone acetyltransferases in Saccharomyces cerevisiae. *Mol. Cell. Biol.*, 26:2791–2802.

[7] Moreira, J. M., S. Holmberg, 1998. Nucleosome structure of the yeast cha1 promoter: analysis of activation-dependent chromatin remodeling of an rna-polymerase-ii-transcribed gene in tbp and rna pol ii mutants defective in vivo in response to acidic activators. *The EMBO Journal*, 17:6028 – 6038.

[8] Gerber, M., A. Shilatifard, 2003. Transcriptional Elongation by RNA Polymerase II and Histone Methylation. *Journal of Biological Chemistry*, 278:26303–26306. doi:10.1074/jbc.R300014200.

[9] Kim, T., S. Buratowski, 2009. Dimethylation of h3k4 by set1 recruits the set3 histone deacetylase complex to 5' transcribed regions. *Cell*, 137:259 – 272. ISSN 0092-8674. doi:DOI: 10.1016/j.cell.2009.02.045.

[10] Wang, A., S. K. Kurdistani, M. Grunstein, 2002. Requirement of Hos2 Histone Deacetylase for Gene Activity in Yeast. *Science*, 298:1412–1414. doi:10.1126/science.1077790.

[11] Santos-Rosa, H., R. Schneider, B. E. Bernstein, N. Karabetsou, A. Morillon, C. Weise, S. L. Schreiber, J. Mellor, T. Kouzarides, 2003. Methylation of histone h3 k4 mediates association of the isw1p atpase with chromatin. *Molecular Cell*, 12:1325 – 1332. ISSN 1097-2765. doi:DOI: 10.1016/S1097-2765(03)00438-6.

[12] Ingvarsdottir, K., C. Edwards, M. G. Lee, J. S. Lee, D. C. Schultz, A. Shilatifard, R. Shiekhattar, S. L. Berger, 2007. Histone H3 K4 Demethylation during Activation and Attenuation of GAL1 Transcription in Saccharomyces cerevisiae . *Mol. Cell. Biol.*, 27:7856–7864. doi:10.1128/MCB.00801-07.

[13] Cinlar, E., 1975. *Introduction to stochastic processes.* Prentice-Hall, Englewood Cliffs, NJ. ISBN 0134980891.

[14] Parzen, E., 1962. *Stochastic Processes.* Holden-Day, San Francisco.

[15] Heyman, D., M. Sobel, 1982. *Stochastic Models in Operations Research.* McGraw-Hill, New York.

[16] Singh, A., J. P. Hespanha, 2009. Optimal feedback strength for noise suppression in autoregulatory gene networks. *Biophys. J.*, 96:4013–4023.

[17] Pedraza, J. M., J. Paulsson, 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*, 319:339–343.

[18] Liu, L., B. Kashyap, G. Templeton, 1990. On the GIX/G/∞ system. *Journal of Applied Probability*, 27:671–683.

[19] Lam, L. T., O. K. Pickeral, A. C. Peng, A. Rosenwald, E. M. Hurt, J. M. Giltnane, L. M. Averett, H. Zhao, R. E. Davis, M. Sathyamoorthy, L. M. Wahl, E. D. Harris, J. A. Mikovits, A. P. Monks, M. G. Hollingshead, E. A. Sausville, L. M. Staudt, 2001. Genomic-scale measurement of mrna turnover and the mechanisms of action of the anti-cancer drug flavopiridol. *Genome Biol*, 2:RESEARCH0041.

[20] Hemberg, M., M. Barahona, 2007. Perfect sampling of the master equation for gene regulatory networks. *Biophys J*, 93:401–10. doi:10.1529/biophysj.106.099390.

[21] Iyer-Biswas, S., F. Hayot, C. Jayaprakash, 2009. Stochasticity of gene products from transcriptional pulsing. *Phys. Rev. E*, 79:9. doi:10.1103/PhysRevE.79.031911.

[22] Zenklusen, D., D. R. Larson, R. H. Singer, 2008. Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.*, 15:1263–1271.

[23] Dobrzynski, M., F. J. Bruggeman, 2009. Elongation dynamics shape bursty transcription and translation. *Proc Natl Acad Sci USA*, 106:2583–8. doi:10.1073/pnas.0803507106.