Science
AAAS

Supplementary Materials for

Stochastic antagonism between two proteins governs a bacterial cell fate switch

Nathan D. Lord, Thomas M. Norman, Ruoshi Yuan, Somenath Bakshi, Richard Losick, Johan Paulsson
Correspondence to: Johan_Paulsson@hms.harvard.edu (J.P.); losick@mcb.harvard.edu (R.L)

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Strain table

Strain/Plasmid Name	Parent	Operation	Genotype
***E. coli* lab stock strains**			
MG1655	n/a	Obtained from CGSC (7740)	*rph-1* λ^-
TB10	W3110	Gift of Dr. Dirk Landgraf	*rph-1* λ^- *nadA::Tn10 λcI857 Δ(cro-bioA)*
PMB14	MG1655	Originally obtained from Dr. Thomas Silhavy	*rph-1* λ^-*attP::Z2(lacI^q tetR)-specR*
JP371	MC4100	Gift of Dr. Dirk Landgraf	MC4100 *rpoS::Tn10 sprE-3xFLAG ΔclpPX::FRT-Kan*
***E. coli* base strains**			
NDL39	MG1655	Transformed with pKD46	*rph-1* λ^-, pKD46
NDL54	NDL39	λ-red integration of ΔmotA allele	*rph-1* λ^- *ΔmotA::FRT-kanR*
NDL56	NDL54	P1 transduction of kanR marker into MG1655	*rph-1* λ^- *ΔmotA::FRT-kanR*
NDL93	NDL56	FLP removal of kanR marker with pCP20	*rph-1* λ^- *ΔmotA*
NDL156	TB10	λ-red integration of P_{Rna1}-mCherry/mKate2-FRT-kanR (henceforth P_{RNA1}-RFP) from pNDL-194	*rph-1* λ^- *nadA::Tn10 λcI857 Δ(cro-bioA) glmS::P_{RNA1}-RFP-FRT-kanR*
NDL159	NDL93	P1 transduction of kanR marker from NDL-156 into NDL93	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP-FRT-kanR*
NDL162	NDL159	FLP removal of kanR marker with pCP20	*rph-1* λ^- *ΔmotA glms::P_{RNA1}-RFP*
***E. coli* spontaneous pulsing reporters**			
NDL339	TB10	λ-red integration of P_{Rv1}-GFPmut2-FRT-kanR	*rph-1* λ^- *nadA::Tn10 λcI857 Δ(cro-bioA) phoA:: P_{Rv1}-GFPmut2-FRT-kanR*
NDL348	NDL162	P1 transduction of NDL339 KanR marker into NDL162	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2-FRT-kanR*
NDL368	NDL348	FLP removal of kanR marker with pCP20	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2*
NDL374	NDL368	P1 transduction of specR marker from PMB14 into NDL368	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2 attP::Z2(lacI^q tetR)-specR*
TMN1221	NDL374	Tn7-mediated integration of pNDL-316 $P_{rpsOmut}$-*SinR*-P_{lac}-*SinI*	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2 attP::Z2(lacI^q tetR)-specR attTn7::* $P_{rpsOmut}$-*sinR*-P_{lac}-*sinI*
NDL423	TMN1221	P1 transduction of *ΔclpPX* allele from JP-371 into TMN1221	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2 attP::Z2(lacI^q tetR)-specR attTn7::* $P_{rpsOmut}$-*SinR*-P_{lac}-*SinI ΔclpPX::FRT-Kan*
NDL406	TB10	λ-red integration of pNDL-305 P_{RV1}-*slr-ssrA-camR*	*rph-1* λ^- *nadA::Tn10 λcI857 Δ(cro-bioA) intC::* P_{RV1}-*slr-camR*
NDL411	TMN1221	P1 transduction of *slr* allele from NDL-406 into TMN1221	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2 attP::Z2(lacI^q tetR)-specR attTn7::* $P_{rpsOmut}$-*SinR*-P_{lac}-*SinI intC::* P_{RV1}-*slr-camR*
NDL419	NDL411	P1 transduction of Δ*clpPX* allele from JP-371 into NDL411	*rph-1* λ^- *ΔmotA glmS::P_{RNA1}-RFP phoA:: P_{Rv1}-GFPmut2 attP::Z2(lacI^q tetR)-specR attTn7::* $P_{rpsOmut}$-*SinR*-P_{lac}-*SinI intC::* P_{RV1}-*slr-camR ΔclpPX::FRT-Kan*

NDL425	NDL162	P1 transduction of *ΔclpPX* allele from JP-371, SpecR marker from PMB14 and integration of a Tn7 construct containing an IPTG-inducible GFP into NDL-162	*rph-1 λ⁻ ΔmotA glmS::P_{RNA1}-RFP attP::Z2(lacI^q tetR)-specR attTn7::* P_{Lac}*-GFP-intC::* P_{RV1}*-slr-camR ΔclpPX::FRT-Kan*
***B. subtilis* lab stock strains**			
3610	n/a	n/a	Wild *B. subtilis* isolate NCIB3610
PY79	n/a	n/a	Prototrophic lab strain PY79
RL4553	3610	n/a	*slr::tet*
RL4928	PY79	n/a	*ylnF/yloA::Tn917::amyE::cat*
RL4930	PY79	n/a	*ywrK::Tn917::amyE::cat*
***B. subtilis* base strains**			
TMN1114	PY79	From (*9*)	P_{tapA}*-cfp* (Cm)
TMN690	3610	From (*9*)	*amyE::* P_{hag}*-gfp (Cm) sacA::*P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)*
TMN694	3610	From (*9*)	*amyE::* P_{hag}*-gfp (Cm) sacA::*P_{tapA}*-mKate2L (Kan) hagA233V (Phleo) slr::mls*
***B. subtilis* HALO reporter strains**			
TMN1033	PY79	Transformed with pTMN1231 and selected for integration at alternative *amyE* locus integrated between *ylnF* and *yloA*.	*ylnf/yloA::*P_{sinI} *-sinI-HALO(spc)*
TMN1107	PY79	Transformed with pTMN1232 and selected for integration at alternative *amyE* locus integrated between *ylnF* and *yloA*.	*ywrK::*P_{sinR} *-sinR-HALO*(spc)
TMN1075	3610	Markerless deletion of *sinI* (see below for details)	*ΔsinI*
TMN1041	TMN1075	*sinI-HALO* allele transduced from TMN-1033 into TMN1075	*ΔsinI ylnF/yloA::*P_{sinI} *-sinI-HALO (spc)*
TMN1050	3610	Long flanking homology deletion (see below)	*sinR::mls*
TMN1123	TMN1041	*tapA* reporter transduced from TMN-1114 into TMN1041	*ΔsinI ylnF/yloA::*P_{sinI}*-sinI-HALO* (spc) P_{tapA}*-cfp*(Cm)
TMN1125	TMN1050	*sinR-HALO* allele transduced from TMN-1107 to TMN1050	*sinR::mls ywrK::*P_{sinR}*-sinR-HALO* (spc)
TMN1126	TMN1125	*tapA* reporter transduced from TMN-1114 to TMN1125	*sinR::mls ywrK::*P_{sinR} *-sinR-HALO* (spc) P_{tapA}*-cfp*(Cm)
TMN1227	TMN1123	Transduced *slr*::tet allele from RL4553 into TMN1123	*ΔsinI ylnF/yloA::*P_{sinI}*-sinI-HALO* (spc) P_{tapA}*-cfp*(Cm) *slr::tet*
TMN1229	TMN1126	Transduced *slr*::tet allele from RL4553 into TMN1126	*sinR::mls ywrK::*P_{sinR} *-sinR-HALO* (spc) P_{tapA}*-cfp*(Cm) *slr::tet*
***B. subtilis* spontaneous pulsing reporter strains**			
TMN1170	TMN690	Δ*sinR* markerless deletion transduced into TMN690	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinR*
TMN1153	PY79	Transformed with plasmid pTMN1231 and selected for integration at alternative *amyE* locus integrated at *ywrK*	*ywrK::*P_{spank}*-sinR*(Spc)
TMN1178	TMN1170	Inducible *sinR* allele transduced from TMN1153 into TMN1170	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinR ywrK::*P_{spank}*-sinR*(Spc)
TMN1182	TMN1178	*slr*::*mls* allele transduced from TMN694 into TMN1178	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinR ywrK::*P_{spank}*-sinR*(Spc) Δ*slr*

TMN1142	TMN690	Δ*sinI* markerless deletion transduced into TMN690	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinI*
TMN1147	PY79	Transformed with plasmid pTMN1229 and selected for integration at alternative *amyE* locus integrated at *ywrK*	*ylnF::*P_{spank}*-sinI* (Spc)
TMN1150	TMN1142	*slr::mls* allele transduced from TMN694 into TMN1142	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinI slr::mls*
TMN1159	TMN1150	TMN1150 transduced with inducible *sinI* allele from TMN1147.	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinI slr::mls ylnF::*P_{spank}*-sinI* (Spc)
NDL-88	TMN1142	TMN1142 transduced with inducible *sinI* allele from TMN1147.	*amyE::* P_{hag}*-gfp (Cm) sacA::* P_{tapA}*-mKate2L (Kan) hagA233V (Phleo)* Δ*sinI ylnF::*P_{spank}*-sinI* (Spc)
***E. coli* plasmids**			
pNDL-194	n/a	Described below.	Modified ColE1 RNA1 promoter driving constitutive expression of mCherry-mKate2 segmentation marker
pNDL-271	n/a	Described below.	SinR-repressible λ P_L-based synthetic promoter ($P_{synthR1}$) driving GFP expression
pTMN1215	n/a	Described below.	Constitutive *sinR* driven by P_{rpsO}, IPTG-inducible expression construct. Set up for attTn7 chromosomal integration.
pNDL-316	n/a	Described below.	Constitutive *sinR* driven by randomly mutagenized P_{rpsO} promoter, IPTG-inducible *sinI* expression construct. Set up for attTn7 chromosomal integration.
pNDL-302	n/a	Described below.	$P_{synthR1}$*-slr-GFPmut2* expression construct. Mutagenized to tune *slr* expression level.
pNDL-305	n/a	Described below.	Mutated $P_{synthR1}$ driving expression of *slr-ssrA* cassette
***B. subtilis* plasmids**			
pTMN1229	n/a	Described below.	IPTG-inducible *sinI* (P_{spank}-*sinI*)
pTMN1231	n/a	Described below.	IPTG-inducible *sinR* (P_{spank}-*sinR*)
pTMN1231	n/a	Described below.	C-terminal fusion of HaloTag to SinI under native regulation (P_{sinI}-*sinI*-HALO)
pTMN1232	n/a	Described below.	C-terminal fusion of HaloTag to SinR under native regulation (P_{sinR}-*sinI*-HALO)

Plasmid and strain construction

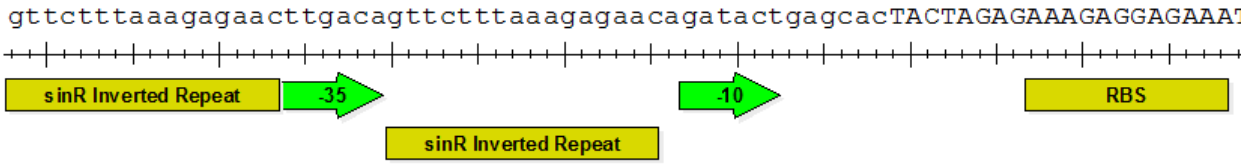
pNDL-194

pNDL-194 links a constitutive RFP expression construct to an FRT-flanked kanamycin resistance maker. This construct encodes the constitutive segmentation marker used in all described *E. coli* strains. The RFP gene encodes an mCherry/mKate2 hybrid protein with the first 11 amino acids of mCherry fused to the full mKate2 ORF (appended sequence: 5' atg gtt agt aaa gga gaa gaa aat aac atg gca 3'). In our hands, this leader appreciably boosts the observed RFP signal in *E. coli*. Expression is driven by a modified constitutive RNA1 promoter from the plasmid colE1 (P_{RNA1}). To generate pNDL-194, the *mKate2* ORF was PCR amplified from pDHL503 (gift of Dr. Dirk Landgraf) with NL-110 and NL-111. Primer NL-110 appends the *mCherry* leader to *mKate2*. Both

primers contained 40 bp of homology to pNDL-52, a plasmid that allows cloning of an insert between the P_{RNA1} promoter and an FRT-kanR allele. The pNDL-52 backbone was PCR amplified with NL-105 and NL-106, and joined with the amplified *mKate2* expression cassette by isothermal assembly (ITA) to yield pNDL-194.

pNDL-271

In order to visualize the state of the reconstituted SinI/SinR circuit, we needed a promoter that responded to SinR expression level. However, as SinR does not exist in *E. coli*, we needed to design this promoter from scratch. Taking inspiration from IPTG-inducible variants of the λ P_L promoter, we placed consensus SinR binding sites between the -10 and -35 elements and immediately upstream of the –35 element of the λ P_L promoter (see diagram below). Notably, this orientation preserves the natural spacing of SinR binding sites in the *B. subtilis epsA-O* operon promoter. This synthetic promoter was synthesized as a gBlock ($P_{synthR1}$ in primer table, IDT) with homology to pNDL-263 (a vector carrying GFPmut2 linked to an FRT-flanked kanamycin resistance marker). Backbone of pNDL-263 was amplified by PCR, and joined with the gBlock by isothermal assembly.



pNDL-302

pNDL-302 encodes the *slr* gene under the control of the synthetic SinR-repressible promoter. We fine-tuned the strength of this promoter by generating a library of pNDL-302 clones with mutated—but still SinR-repressible—promoter sequences. See below for details on the screen design and selection criteria. To generate pNDL-302, a gBlock encoding a $P_{synthR1}$-*slr-ssrA* expression construct (P_{sythR1}-*slr-ssrA* in primer table) was ordered (IDT), and PCR-amplified with NL-284 and NL-287. Both primers appended homology to pUA66 (a vector encoding a promoterless RBS-GFP cassette (*43*)) to targeting insertion upstream of the GFP ORF, and NL-284 contained a series of randomized bases over the -35 element of the promoter. A library of P_{sythR1} variants was generated by fusing the PCR-amplified gBlock to the pUA66 backbone by ITA. Note that the plasmid product was transformed into MC1061, a common strain that (1) has functional clpP-mediated proteolysis to ensure that Slr levels were kept low through active proteolysis, and (2) has no *sinR* gene, ensuring that the cassette was expressed to its maximal level.

pNDL-305

pNDL-305 encodes the finalized $P_{synthR1mut}$-*slr-ssrA* construct linked to a chloramphenicol resistance cassette. To generate pNDL-305, the backbone of pNDL-280 (a pUC-based cloning vector with an MCS upstream of a chloramphenicol resistance gene) was PCR amplified and the $P_{synthR1mut}$-*slr-ssrA* cassette was amplified from the desired pNDL-302 clone. Both fragments were purified and joined by ITA.

pTMN1215

pTMN1215 is a Tn7 integration vector for a cassette containing a constitutively-expressed sinR gene under the constitutive P_{rpsO} promoter and an IPTG-inducible sinI gene. This plasmid serves as the template for the creation of the pNDL-316 mutagenesis library described below.

To construct pTMN1215, we first synthesized a cassette containing a promoterless sinR allele and a synthetic terminator. This construct was then amplified using primer pair PTMN255P1F/R. In separate work, a P_{tet}-sinR-T1 terminator-P_{lac}-sinI-T1 terminator cassette was synthesized and integrated into the plasmid pNDL-281. We amplified the P_{lac}-sinI-T1 terminator portion of this cassette from pNDL-281 using primer pair PTMN255P2F/R. The two products were then ligated by ITA into pUC19.

The product above lacks a promoter for sinR. This was purposefully avoided because we had noted toxicity when sinR was expressed from high copy vectors. Using the Alon fluorescent plasmid library (*43*), we identified three candidate constitutive promoters: P_{rpsO}, P_{rpsU}, and P_{oppA}. Because the Alon plasmid library vectors share a common backbone, we could amplify each of these candidate promoters using the primer pair PTMN256P1F/R. We separately amplified the sinR-synthetic terminator-P_{lac}-sinI-T1 terminator cassette created above using the primer pair PTMN256P2F/R. The products were then ligated by ITA into the low copy vector Tn7 integration vector pGRG36 cut with AscI/PacI, producing pTMN1215. The promise of the candidate plasmids was then tested by transforming them into NDL-374, which contains the synthetic SinR-repressible fluorescent reporter. Based either on the identification of errors during sequencing or lack of repression seen in this assay, it was ultimately decided to proceed with P_{rpsO} as the basis for further mutagenesis.

pNDL-316

pNDL-316 is the Tn7 integration vector for a cassette containing a constitutively-expressed *sinR* gene and an IPTG-inducible *sinI* gene. It was derived from pTMN1215 by random mutagenesis of the *rpsO* promoter. This mutagenesis was performed in order to find an expression regime in which SinI and SinR were produced at comparable levels Several independent pNDL-316 clones were generated and screened to identify variants with matched *sinI*-*sinR* expression levels. See below for a description of the screen design and selection criteria.

To construct pNDL-316, the P_{rpsO}-*sinR*-P_{lac} cassette was amplified from pTMN1215 by PCR with NL-294 and NL-295. Primer NL-294 contains four randomized bases covering part of the -35 sequences of the *rpsO* promoter, meaning that the products will comprise a library of promoter variants. We then PCR amplified the remainder of the pTMN1215 backbone using NL-296 and NL-297, and joined the two products by ITA.

TMN1050

A disruption of the *sinR* locus with an erythromycin resistance cassette was performed using the long flanking homology technique and primers PTMN106P1-P4. Briefly, the flanks of the gene were first amplified using pairs P1/P2 and P3/P4. The interior primers P2/P3 introduce homologous tails that allow the amplified flanks to serve as `mega-primers' for the amplification of an antibiotic resistance cassette. The resulting linear product was transformed directly into

PY79, and resistant transformants were screened for loss of *sinR*. The marked deletion was then carried into 3610 by phage transduction.

pTMN1229/1231

These are the IPTG-inducible P_{spank}-*sinI*/P_{spank}-*sinR* plasmids. (The P_{spank} promoter is a single nucleotide mutation of the $P_{hyperspank}$ inducible promoter that has substantially lowered expression and tight repression by the Lac repressor.) The genes for *sinI* or *sinR* were amplified using primers PTMN237F/R and PTMN236F/R, respectively, and ligated by ITA into pDR110 cut with HindIII and SphI.

Markerless *sinI* deletions

These were performed as described previously (*9*) using phage produced from a strain transformed with pTMN994. Since there is an internal *sinR* promoter within the *sinI* ORF, this deletion removes only the ribosome binding site and the first 57 nucleotides.

pTMN1231/1232

These plasmids contain the translational fusions of the HaloTag to *sinI* and *sinR* under their native regulation, which we denote as P_{sinI}-*sinI*-HALO and P_{sinR}-*sinR*-HALO.
To generate pTMN1231, a large upstream fragment upstream was amplified alongside *sinI* using the primer pair PTMN207F/P2-I. This fragment includes a putative upstream promoter within the *yqhG* ORF identified by the *B. subtilis* Expression Data Browser. The HaloTag was then amplified from a plasmid provided by Dirk Landgraf using PTMN207P3-I/R. The two fragments were then joined in a 3-part ITA reaction into the plasmid backbone pDG1730 cut with BamHI and EcoRI.
There is an internal *sinR* promoter within the *sinI* gene that we wanted to preserve while avoiding introducing an additional copy of *sinI*. Thus to produce pTMN1232 we amplified a large upstream fragment (again containing the putative upstream promoter within *yqhG*) that ran to the end of the *sinI* ORF using primer pair PTMN207F/P2-R, but we used genomic DNA from strain TMN1075 as a template. This strain contains the markerless deletion of *sinI* described above that only the 5' end of the ORF and leaves the *sinR* promoter intact. Like in pTMN1231, the HaloTag was then amplified using PTMN207P3-R/R and the two fragments were ligated into pDG1730.

Screening for $P_{synthR1}$ variants in pNDL-302

The Slr protein induces substantial toxicity in *E. coli* when expressed to even moderate levels. When cloned without a synthetic degradation tag, we routinely recovered alleles with nonsense mutations. In an effort to reduce this toxicity, we targeted Slr protein for clpP-mediated degradation by appending an ssrA tag to the C-terminus. To further tune the expression level, we generated a library of P_{sythR1} promoter variants by randomly mutagenizing the -35 element (see details on construction of pNDL-302 above).
When designing the screen, we were concerned that most of the clones we recovered would simply have no promoter activity. To screen out such variants, we appended an RBS-GFPmut2 cassette to the end of the *slr-ssrA* gene, thereby creating a synthetic operon. By selecting only fluorescent colonies, we could identify clones with appreciable mRNA production. We picked several independent clones with a range of colony GFP intensities and sequenced the *slr* gene in each

clone. The clone used for subsequent *slr* manipulations carried two substitutions in the promoter's -35 site (TTGACA→ TGTACA). We refer to this promoter as $P_{synthR1mut}$ from this point on.

Screening of P_{rpsO} variants in pNDL-316

The central tasks in reconstituting the SinI-SinR system in *E. coli* were to identify synthetic promoters that allowed (1) *sinR* to be expressed constitutively, (2) *sinI* to be expressed inducibly (but stably in time), and (3) for both proteins to be expressed at similar levels. For inducible *sinI* expression we tried both IPTG and aTc-inducible promoters. In our hands, aTc rapidly degraded in mother machine experiments, making it impossible to achieve stationary switching statistics. IPTG, on the other hand, appeared stable over several days at 37 C as evidenced by the fact that mean GFP expression remained stable over this period. We thus chose to use P_{Lac} to control *sinI* expression. For *sinR*, we therefore had to find or generate a constitutive promoter with output comparable to the strength of our IPTG-inducible promoter at 'mid-saturation' levels of inducer.

We chose to generate and screen a small library of variants of the promoter for the ribosomal gene *rpsO*. See cloning above for details of the construction of the library (pNDL-316). Each variant in the library differed in its -35 promoter element. To identify clones of interest, we transformed the library into NDL-374, which contains both a $lacI^q$ allele, a SinR-repressible $P_{synthR1}$-GFP reporter gene, and an IPTG-inducible *sinI* gene, and replica-plated a small library of clones on $IPTG^+$ and $IPTG^-$ (1 mM IPTG and 0 mM IPTG, respectively) medium. We then isolated clones that, as judged on the fluorescent macroscope, yielded colonies that were GFP positive in the presence of IPTG (when *sinI* was maximally expressed) and GFP negative in the absence of IPTG (when *sinI* expression was repressed by LacI). This selection criterion guarantees that the range of available *sinI* expression levels covers the full range of stochastic competition behaviors: from complete dominance by SinR to complete dominance by SinI. The clone used to generate TMN-1221 contained both a point mutation and an insertion TGCG-ATAA → TGAGTATAA.

Rationale behind strain background of *E. coli* reconstitution

The strain background used for reconstituting the SinIR system in *E. coli* has several conspicuous features.

(1) Deletion of the *motA* gene. Motile *E. coli* readily swim out of the mother machine channels, making long-term imaging nearly impossible in many wild type strains. To paralyze our background strain, we knock out one component of the flagellar motor, *motA.* In *motA* null strains, fully functional flagella are still constructed, however they cannot generate force, leaving the cells largely immobile.

(2) Inclusion of a constitutive RFP expression cassette. Transmitted light images in the mother machine are of exceedingly low quality, and are therefore difficult to use for automated cell segmentation. To facilitate image analysis, we express a bright, constitutive RFP reporter in all strains.

(3) Deletion of the *clpP* and *clpX* genes. This manipulation removes the ClpXP protease, which is responsible for degradation of ssrA-tagged substrates. As described above, we actually ssrA-tag

the *slr* gene to reduce its expression. When we first monitored the pulsing of a strain bearing a chromosomally-integrated *slr-ssrA gene* in addition to the reconstituted SinI-SinR circuit, we saw no difference in pulsing behavior relative to the SinI-SinR circuit alone (data not shown). We interpreted this result as meaning that we had driven the *slr* expression level down too far to see an effect on pulsing. To increase *slr* levels, we then transduced a *clpPX* null allele into both the SinI-SinR and the SinI-SinR-SlR circuit. These two strains are compared in Figure 3 of the main text. We note that pulsing does occur in a *clpPX*-positive background (Movie S4): we chose this approach because it facilitates integration of *slr* at single copy without passing through a high-expression intermediate, and therefore reduces the risk of nonsense mutations. All cloning of the *slr-ssrA* cassettes was done in strains with functional ClpXP protease—which kept *slr* levels low enough to avoid toxicity—and deletion of *clpX* and *clpP* allowed us to dial up expression again once the gene was present only at single copy.

Primer table

Name	Sequence
PTMN106P1	AAGGCTTAGGCTATATGACAGTGCAGC
PTMN106P2	caattcgccctatagtgagtcgt GATATTATAGCACATTCAGAAAGGATTTACGG
PTMN106P3	ccagcttttgttccctttagtgag AAAAATGGAGAAAATCCCAAAAAGAGG
PTMN106P4	CAGCGCCATTAGAGAAATTGAAAGAAAG
PTMN236F	TGTGTGGAATTGTGAGCGGATAACAATTA AGCTT AAGGAAGGTGATGACATTGATTGGC
PTMN236R	TGATGACCTCGTTTCCACCGAATTAGCTTG GCTCAGGCACTACTCCTCTTTTTGGG
PTMN237F	TGTGTGGAATTGTGAGCGGATAACAATTA AGCTT AGGAGGAGAAACTGCATGAAGAATG
PTMN237R	ATGACCTCGTTTCCACCGAATTAGCTTG GATATTATAGCACATTCAGAAAGGATTTACGG
PTMN255P1-F	acgttgtaaaacgacggccagtg ttgattggccagcgtattaaac
PTMN255P1-R	GTTATCCGCTCACAATTCTCGAG GGAGCGGATACATATTTGAATGCC
PTMN255P2-F	GGCATTCAAATATGTATCCGCTC CCTCGAGAATTGTGAGCGGATAAC
PTMN255P2-R	gctatgaccatgattacgccaagct GCTGCGTTCGGTCAAGGTTCTG
PTMN256P1-F	TATCGGATCCTAGTAAGCCACGTTTTA GCGTATCACGAGGCCCTTTC
PTMN256P1	GTTTAATACGCTGGCCAATCAA ATGTATATCTCCTTCTTAAATCTAGAGGATCC
PTMN256P2-F	GGATCCTCTAGATTTAAGAAGGAGATATACAT TTGATTGGCCAGCGTATTAAAC
PTMN256P2-R	TCGACGCGGCCGTGG GCTGCGTTCGGTCAAGGTTCTG
PTMN207F	*AAACACACAAATTAAAAACTGGTCTGATCG* GCGCCAAAAGACCTAGATGGTGA
PTMN207P2-I	*TACTGCCACCGCCACCGCT* GAAAGGATTTACGGTATGACTTCTGGC
PTMN207P2-R	TACTGCCACCGCCACCGCT CTCCTCTTTTTGGGATTTTCTCCATT
PTMN207P3-I	GCCAGAAGTCATACCGTAAATCCTTTCAGCGGTGGCGGTGGCAGTA
PTMN207P3-R	*AATGGAGAAAATCCCAAAAAGAGGAG* AGCGGTGGCGGTGGCAGTA
PTMN207R	*ATTCGCCAGGGCTGCAGG* TTAACCGGAAATCTCCAGAGTAGACAGC
NL-105	GGA TCC ATC TCC TTC TTA AAT GAA TTC AAA TAC
NL-106	GCT TAA TTA GCT GAG TCT AGA GGC ATC
NL-110	AAG GAC AGT ATT TGA ATT CAT TTA AGA AGG AGA TGG ATC CAT GGT TAG TAA AGG AGA AGA AAA TAA CAT GGC ACT GAT TAA GGA GAA CAT GCA CAT GAA GC
NL-111	TTT CGT TTT ATT TGA TGC CTC TAG ACT CAG CTA ATT AAG CTT ATC TGT GCC CCA GTT TGC
NL-284	GAG GCC CTT TCG TCT TCA CCT CGA GGT TCT TTA AAG AGA ACT NNN NAG TTC TTT AAA GAG AAC AGA TAC TGA GC
NL-287	ATC TCC TTC TTA AAT CTA GAG GAT CTC AAG CTG CTA AAG CGT AGT TTT CGT CGT TTG CTG CTC TTC CCT TTG TTT TTA AAA AGG
NL-294	AAT ACC CGG CGT AAT GTT AAC CGT CTN NNN ATA ACA GGT CGC TAC G
NL-295	CTT GTT TAA ATA AGC TCG CCA TAT CCA CTA TTA CCC C
NL-296	ACG GTT AAC ATT ACG CCG GGT ATT CAA CC
NL-297	GGG GTA ATA GTG GAT ATG GCG AGC TTA TTT AAA CAA G

Mother Machine Imaging Protocol

Chip preparation

PDMS devices were prepared as previously described (*9*). Briefly, Sylgard 184 monomer was mixed in a 5:1 ratio with curing agent (Dow Corning), poured onto the appropriate silicon master, degassed under vacuum, and baked at 65 C for at least 1 hour. Individual devices were cut from the wafer, inlets and outlets were inserted using biopsy punches, and the device was bonded to a dry, KOH-cleaned No. 1.5 coverglass by treatment with oxygen plasma (30 second treatment at 50 W with 170 mTorr O_2 pressure). Bonded devices were cured for an additional hour at 65 C.

Cell growth conditions and device loading

For *B. subtilis* experiments, cells were prepared and loaded into microfluidic devices as described in (*9*). For *E. coli* experiments, strains were revived from glycerol stocks by streaking and growing overnight at 37 C on LB plates containing appropriate antibiotics. The next day, 5-mL LB cultures were started and shaken overnight. In the morning, overnight cultures were diluted 1:300-1:500, shaken to OD_{600}=0.3-0.6, and injected into the feeding channel of a microfluidic device. The chip was then mounted on a custom-machined adapter that allowed it to be mounted into a benchtop microcentrifuge and spun at $\geq$ 5000 *g* for 10 minutes. Medium flow was driven using 60-mL syringes mounted on syringe pumps (set to 5-7 uL/minute, New Era Pump Systems). Syringes were connected to each feeding channel through narrow gauge Tygon tubing (VWR). In *E. coli* experiments, we supplemented all LB medium with Pluronic F108 (Sigma-Aldrich) to a concentration of 0.65 g/L. This surfactant dramatically reduced cell adhesion to the walls of the feeding channel, thereby extending possible duration of experiments. In experiments requiring it, IPTG was first introduced to the cells after loading by including it in the growth medium. We note that the exceptionally-long time traces afforded by the mother machine devices allow us to directly observe the period of adjustment to new conditions (the device environment and presence of IPTG, typically lasting 5-10 generations), and remove this transient from the dataset.

Live cell microscopy

Image acquisition was performed as previously described (*9*). Briefly, microscopy was carried out using a Nikon Eclipse Ti invertied microsope with a 1.4 NA 60X Plan Apo oil immersion objective (Nikon), automated stage (Ludl), custom-built incubator (temperatures were maintained at 37 C), LED illumination system (Lumencor SOLA Light Engine), and an Orca R2 CCD camera (Hamamatsu). All acquisitions were carried out using custom-made MATLAB scripts executed through μManager [RW.ERROR - Unable to find reference:269]. Filter sets used for imaging were: Semrock GFP-1828A, Semrock mCherry-B, Semrock CFP-2432C, Semrock YFP-2427B, and Semrock Cy3-4040C (for TMR). In *E. coli* experiments, segmentation images (RFP channel) were taken without binning, and all other images were 2x2-binned to reduce exposure times. For *B. subtilis* experiments, all images were taken with 2x2 binning. Focus was maintained over long-term experiments using a custom autofocus algorithm (operating on a dedicated imaging position that was not used for data analysis) and a Nikon PerfectFocus system. All images were saved as 16 bit TIFFs. After loading, cells were allowed to equilibrate to the device conditions for several hours before imaging began. For both organisms, images were collected every 8-10 minutes, yielding 2-3 images per cell generation.

Microfluidic Master Fabrication

Fabrication of templates for all microfluidic devices were carried out using conventional UV photolithography in a clean room environment. The device was designed using AutoCad, and quartz-chrome photomasks were ordered from Toppan Inc. and The Center for Nanoscale Systems at Harvard University. The *B. subtilis* 'two-tiered' design was previously described (*9*), and details of its construction are contained in the supplementary methods of that paper. The *E. coli* device was fabricated in a very similar fashion, and the steps are included below. AutoCad files for our *E. coli* device are similar to those described by Wang *et al.* (*25*), but are available on request.
As in our previous paper, we note that the process parameters described below should be regarded as starting points as the tolerances on the cell channels appear to be close to the achievable resolution of UV photolithography. We therefore recommend that, when following our protocol, that several variant masters are made in parallel, all with slightly different fabrication details. This ensures that, in the end, at least one useful device template is obtained. Note: all spin steps below use the shorthand notation speed (rpm)/acceleration (rpm/sec)/time (seconds).

First Layer: Su8 Base Coat

In our hands, the key problem in fabricating this device is in achieving high adhesion of the very small cell channels. Su8 adhesion to silicon is good but not perfect, and we found that these channels frequently distorted or delaminated during development. To combat this, we first spin and cure a thin, unpatterned layer of Su8 onto the wafer to serve as an 'adhesive'.

1) Place a new 3" Si wafer (we used 380 um TEST grade wafers from University Wafer) in a dish of fresh acetone. Sonicate at high power for 5 minutes.
2) Sequentially rinse the wafer with streams of methyl alcohol (MeOH), isopropyl alcohol (IPA) and H_2O (~ 10 seconds per solvent).
3) Place wafer on 2" spin chuck and spin seconds at 500 rpm.
4) While spinning, sequentially rinse the wafer with streams of MeOH, IPA and H_2O.
5) Spin wafer 1 minute at 3000 rpm to dry.
6) Dehydrate wafer 15 minutes on a hot plate set to 150°-200° C.
7) Set spin program to: Step 1: 500/100/10, Step 2: 3000/300/60.
8) Place the dehydrated wafer onto the spin coater chuck and dispense a small (cover ~2/3 of the wafer surface) amount of Su8 2000.5 photoresist (Microchem) with a pipette. Run the spin program. This should result in a coat of ~0.5 um.
9) Soft bake wafer (in order) for 1 minute at 65° C, 1 minute at 95° C, 1 minute at 65 C.
10) Expose wafer 5 seconds with no mask at 25 mW/cm^2 (I-line).
11) Post exposure bake the wafer (in order) 1 minute at 65° C, 1 minute at 95° C, 1 minute at 65° C.
12) Do not develop or hard bake the wafer. Instead, proceed directly to the second layer protocol.

Second Layer: Cell Channels

This set of steps lays down the channels that house the cells in the final device. The tolerances for this layer are very stringent; the exposure dose and contact between mask and wafer must be optimized. We recommend trying a range of exposure parameters to ensure that a useful device is obtained. We also stress the importance of the very long post exposure bake time in the process below. In our hands, this greatly improves the retention of the cell channels during development.

1) Set spin program to: Step1: 500/100/10, Step 2: 3500/300/60.

2) Place the wafer onto the spin coater chuck and dispense a small (cover ~2/3 of the wafer surface) amount of Su8 2002 photoresist with a pipette. Run the spin program. This should result in a coat of ~1.3 μm.
3) Soft bake wafer (in order) 1 minute at 65° C, 3 minutes at 95° C, 1 minute at 65° C.
4) Expose wafer 0.75 seconds (25 mW/cm^2, I-line) through cell channel mask in vacuum contact mode.
5) Bake wafer 1 minute at 65° C, 20 minutes at 95° C, 1 minute at 65° C.
6) Develop wafer 30 seconds with very gentle agitation in Su8 developer.
7) Rinse wafer 10 seconds with IPA.
8) Check completeness of development process. If undeveloped Su8 remains on the wafer (other than the desired cell channels) repeat developer treatment for 10 seconds.
9) Hard bake wafer 10 minutes at 150° C.
10) Verify channel height using a profilometer. The expected height is 1.5 μm. If the channels dimensions lie outside of your expected tolerance bounds, the process must be repeated with modified spin coating parameters.

Third Layer: Feeding Channels

The final layer of the device forms the medium flow channels. The dimensions of these features are not critical: We have used feeding channels of widely varying dimension to similar effect. The alignment is sensitive to large errors, however. The alignment between feeding channels and cell channels must be accurate (down to a couple of microns) in order to ensure that the cell channels are of the desired final length.

1) Set spin program to: Step 1: 500/100/10, Step 2: 5000/300/60.
2) Place the wafer onto the spin coater chuck and dispense a small (cover ~2/3 of the wafer surface) amount of Su8 2025 photoresist with a pipette being careful not to introduce bubbles. Run the spin program. This should result in a coat of ~15 um.
3) Soft bake the wafer (in order) 1 minute at 65° C, 4 minutes at 95° C, 1 minute at 65° C.
4) With an Su8-developer-soaked swab, clean the newly-deposited photoresist off of the alignment marks to make them visible for the alignment process.
5) Soft bake the wafer (in order) 1 minute at 65° C, 4 minutes at 95° C, 1 minute at 65° C.
6) Align feeding channel mask to the alignment marks on the wafer. Apply vacuum contact and check alignment again. If the vacuum application skewed the alignment, repeat the alignment process.
7) Expose wafer 10 seconds (25 mW/cm^2, I-line) through aligned feeding channel mask.
8) Bake wafer 1 minute at 65° C, 4 minutes at 95° C, 1 minute at 65° C.
9) Develop wafer 1.5 minutes in Su8 Developer with mild agitation.
10) Rinse wafer 10 seconds in IPA. Check to ensure that the development is finished. If undesired photoresist remains, develop again for 20 seconds.
11) Hard bake wafer 15 minutes at 150 C.
12) Verify channel height using a profilometer. The expected height is 15 μm. If the channels dimensions lie outside of your expected tolerance bounds, the process must be repeated with modified spin coating parameters.

Image Analysis

Cell segmentation

Image segmentation was carried out using custom MATLAB-based scripts that were described previously (*9*). Briefly, *E. coli* movies were segmented using the signal from a constitutively-expressed mKate2 construct. *B. subtilis* movies were segmented using the motility reporter signal. We note that, because *slr* was deleted in all *B. subtilis* experiments reporter reported here, firing of the SinI-SinR circuit did not result in repression of the motility regulon. As a result, we could safely rely on this signal to segment all cells. Fluorescence images were contrast-enhanced using an unsharp mask filter, and transformed into binary images using a standard edge finding algorithm (a 'Laplacian of Gaussian' filter). Merged cells were then split in two phases. First, the fluorescence intensity along the midline of each cell was searched for local minima. These sites marked nascent cell septa, and were cut after identification. Second, points of constriction in each object were identified and cut. Finally, the boundaries of all identified cells were refined by morphological opening and thickening. Spurious objects (i.e. objects with abnormal eccentricity, length or area) were removed from the segmentation.

Fluorescence measurements

Before recording pixel values within each cell, fluorescence background was removed from each raw image by subtracting the median pixel intensity. Fluorescence intensities for each identified cell were measured both by the mean pixel intensity of all object pixels and as the median of the brightest 10% of pixels within each object. Analysis on both of these quantities yields nearly identical results. We therefore report results only for the mean intensity measurements.

Lineage compilation

Lineage tracking was performed as previously described (*9*). In brief, we track only the 'mother cell' in each channel (i.e. the cell closest to the 'dead end' of each cell channel), as this is the only cell that remains in the device throughout the experiment. This practice is equivalent to following one particular branch of the lineage tree. Mother cells were matched between frames by simple centroid matching: as the mother cell remains essentially stationary in the x-coordinate throughout the movies, interframe matches can be inferred by matching cells with the closest x-coordinates (within a previously described acceptable range). Cell divisions were identified abrupt decreases in cell area between frames. All lineages were manually curated to remove spurious segmentation errors and dead cells that were connected into lineages.

We note that, in both *E. coli* and *B. subtilis*, a transient adaptation period—due to adjustment to the device environment and presence of IPTG in the medium—of 5-10 generations was often observed. To deal with this issue, we leverage the exceptional length of lineage traces generated by the mother machine. We empirically identify the spans over which the cell growth and reporter output statistics are stationary (i.e. have a stable mean and variance in time), and use only this data. Because our traces are often more than 100 generations long, removal of the 5-10 generation transient at the beginning of the movies does not significantly reduce sampling.

Flat field correction

Uneven illumination poses a serious impediment to measurement of timescales with autocorrelations. While such effects are often small in magnitude (i.e. ~10% CV with LED illumination), they are static throughout each experiment, leading to autocorrelation curves with an 'offset' that never decays away. This offset disrupts the estimation of fluctuation timescales, and makes experimental data near impossible to compare with theoretical expectations. We therefore apply flat field corrections to the raw intensity data prior to calculation of autocorrelation functions.

The constant position of each cell within the mother machine allows for effective correction without the need for an external standard. For each strain and condition, the intensity values for all segmented cells in the experiment are pooled and sorted by x centroid position. A sliding window average of cell intensity as a function of spatial position is then compiled. Because the average intensity at each position would be constant in the absence of uneven illumination, this curve estimates the local illumination intensity at each position. Each observation is then multiplicatively corrected using the appropriate entry of the correction curve. We note that this is a 1-dimensional correction, as the cells are grouped only by x-coordinate. The spatial patterning of cells in the mother machine justifies this assumption: mother cells therefore lie along a line (parallel to the x-axis in our images), and the variation in illumination intensity perpendicular to this line is negligible because of the small size of each cell.

Data Analysis

Summary of mother machine dataset

The table below summarizes the bulk statistics of mother machine data acquisition. We record the total number of independent lineages (i.e mother cells) observed, the total number of divisions of these mother cells, and the total number of SinR reporter 'firing' events observed. Summary statistics for the SinI-HALO and SinR-HALO dataset are also included at the bottom of the table.

Circuit	Strain	Condition	Figure Panels	Recorded Lineages	Mother Cell Division	Switching Events
E. coli Reconstitution (SinI + SinR)	NDL-423	90 μM	Figure 2 (panels H-K) Figure 3 (panel A)	135	13919	560
E. coli Reconstitution (SinI + SinR)	NDL-423	100 μM	Figure 2 (panels H-K) Figure 3 (panel A)	257	25246	1434
E. coli Reconstitution (SinI + SinR)	NDL-423	110 μM	Figure 2 (panels H-K) Figure 3 (panel A)	108	12213	1106
B. subtils Reduced (inducible SinI)	TMN-1159	7.5 μM	Figure 2 (panels D-E)	89	16510	359

B. subtils Reduced (inducible SinI)	TMN-1159	10 μM	Figure 2 (panels D-E)	95	16474	898
B. subtils Reduced (inducible SinI)	TMN-1159	12.5 μM	Figure 2 (panels D-E)	92	5816	498
E. coli Feedback Reconstitution (SinI+SinR+SlrR)	NDL-419	60 μM	Figure 3 (panels A-B)	61	4387	239
E. coli Feedback Reconstitution (SinI+SinR+SlrR)	NDL-419	80 μM	Figure 3 (panels A-B)	69	6049	467
E. coli Feedback Reconstitution (SinI+SinR+SlrR)	NDL-419	140 μM	Figure 3 (panels A-B)	63	4085	265
B. subtilis Native	TMN-1157	LB only	Figure 1 (panels E-F)	74	20792	470
SinI-HALO	TMN-1227	LB only	Figure 1 (panel B)	5289 cells total	n/a	n/a
SinR-HALO	TMN-1229	LB only	Figure 2 (panel B)	8736 cells total	n/a	n/a

Peak finding

We briefly summarize the computational procedure used to identify gene expression pulse boundaries. For a full description, please refer to (*9*).

The rough peak locations were identified by searching for local maxima in the matrix reporter (P_{tapA}-*cfp*, *B. subtilis*), or the synthetic SinR-repressible promoter ($P_{synthR1}$-*gfp*, *E. coli*) using 'peakfinder', an open source MATLAB function (Nathanael C. Yoder, available at http://www.mathworks.com/matlabcentral/fileexchange/25500-peakfinder). On either side of this point, fluorescence intensities decrease rapidly. We identify the boundaries of the peak by looking for the ends of this rapid descent, that is, where fluorescence intensities maintain essentially constant values. We then estimate and subtract the fluorescence baseline for each trace. Peak boundaries were then refined to reflect where the signal surpasses this background. That is, the beginning and end of each peak are set to be the times at which the fluorescence signal unambiguously surpasses and drops below (respectively) this background level. After this refinement, all peaks were manually curated and compared to the raw images to ensure that (1) each event represented a *bona fide*, distinct gene expression pulse, (2) overlapping peaks were split from one another and (3) remaining errors in segmentation were manually removed from the dataset.

We illustrate this process with example data in Fig. S18, and demonstrate that the key statistical properties of the residence time in the ON state are insensitive to choice of background threshold

in Fig. S19. Additionally, we demonstrate that the sub-exponential distribution of pulse durations can be inferred without calling discrete events (see Fig. S20 and 'Independent inference of pulse timing' section, below), and that comparable switching statistics result from calling events based on cessation of expression of a motility reporter (Fig. S21).

Definitions of key quantities

We next define the various quantities measured in our microfluidic experiments. Procedures for simulating theoretical expectations are described in the section *Simulations in Figures 1 and 2 of the main text*.

(1) *Inter-peak time*: This quantity measures the time elapsed between consecutive firings of the SinR-responsive reporters (P_{tapA} in *B. subtilis* and $P_{synthR1}$-*gfp* in *E. coli*). If two adjacent peaks begin at times t_1 and t_2, then the subsequent initiation time for this pair is $t_2 - t_1$. We note that this quantity can be distinguished from the 'motility duration', which we defined in previous work as the time between the end of one pulse and the beginning of the next. Because of the exceptionally long duration of residence times in the OFF state, these statistics are nearly identical, and we report only the inter-peak times.

(2) *Peak duration*: The peak duration measures the length of a gene expression event. As described above, the beginning and end of each gene expression peak are defined as the points at which it crosses a pre-defined threshold, chosen so as to mark the intensity at which a signal is unambiguously distinguishable from fluorescence background. If a peak's leading and trailing edges cross this threshold at times t_1 and t_2, respectively, then the pulse period is defined as $t_2 - t_1$.

(3) *Fraction ON*: In Figure 2, we measure the population fraction of cells in the 'ON' state. For the experimental figures, we simply report the fraction of time the reporter intensity lies above a threshold chosen to be conservatively above the average intensity of cells in medium without IPTG. We separately carried out experiments to ensure that the observed ultrasensitivity arose from the SinI-SinR interaction and not from the cooperativity of our IPTG-inducible promoter; see below.

(4) *GFP Concentration*: We report gene expression levels as GFP concentrations throughout the text. These concentrations are operationally defined as the average pixel intensity within a given cell's segmentation mask. Alternate measures of the concentration (for example, the mean of the brightest decile of pixels) lead to nearly identical results.

(5) *GFP production rate*: We use derivatives in the GFP signal at two points in the paper. (a) To infer properties about the distributions of SinI dominance times. For this purpose, we define the GFP production rate as the total amount of GFP produced between two adjacent frames. For reasons explained below ('Calculating of GFP production rates under saturating [IPTG]'), we only define this quantity on frames that do not span a cell division, and can therefore define it as the difference in total intensity (i.e. sum of all pixel values within the mask) between frames $i+1$ and i. (b) To infer the presence of pulse timing without calling discrete events (i.e. Fig. S20). In this case, we define GFP production rate as the difference in GFP concentration (i.e. average pixel intensity within a cell) between successive frames $i+1$ and i.

Calculation of reporter autocorrelations

The autocorrelation is defined for a stationary signal $x(t)$ with mean and variance $\langle x \rangle$ and σ_x^2 as:

$$A(\tau) = \frac{\langle (x(t) - \langle x \rangle)(x(t+\tau) - \langle x \rangle) \rangle}{\sigma_x^2}$$

with angled brackets denoting averages. While calculating this quantity for a single, long time trace is straightforward—the meanings of the average brackets, mean, and variance are obvious in this case—several different estimators may be used when the data is broken up into multiple, independent time traces.

For a collection of *n* independent time traces, where $x_i(t)$ denotes the value of *i*th trace at time *t*, we estimate the autocorrelation using the following expression:

$$A(\tau) = \frac{\langle \langle (x_i(t) - \langle x_i(t) \rangle_i)(x_i(t+\tau) - \langle x_i(t+\tau) \rangle_i) \rangle_t \rangle_i}{\sigma_x^2}$$

where $\langle \cdot \rangle_i$ denotes averaging over lineages, $\langle \cdot \rangle_t$ denotes averaging over time points, and σ_x^2 denotes the steady state variance in *x*. We estimate σ_x^2 as follows:

$$\sigma_x^2 = \langle \langle (x_i(t) - \langle x_i(t) \rangle_i)^2 \rangle_t \rangle_i$$

We note that, in practice, the population averages ($\langle x_i(t) \rangle_i$) in the above expressions are calculated using a 20-frame sliding window. That is, $\langle x_i(t) \rangle_i$ is the average of all data collected between frames *t* - 10 and *t* + 10. This practice results in a more stable estimate of the population average for finite sampling.

In both of these estimators, the chosen order of averaging (traces first, time second) means that we first estimate the quantity (autocovariance or variance) for each trace individually, then average these estimates together. We chose this approach because, in our hands, it leads to the most reproducible estimates for the segmentation reporters in both *E. coli* and *B. subtilis*. In this sense, we use the segmentation reporters as positive controls, and assume that they should yield the same autocorrelations in all tested conditions.

In some time-lapse experiments, we found that the autocorrelations sometimes decayed to a non-zero offset after ~10 generations. These offsets were generally small (i.e. ~0.1 after normalizing the functions). Such behavior can occur if flat-field correction is imperfect or if inducer concentrations decay slowly over the course of the several days needed for each experiment. In order to fairly compare the correlations of each dataset spanning ~10 generations, we estimated these small offsets, subtracted them and renormalized the autocorrelation functions.

We derived error bars for the autocorrelation functions by bootstrapping (see e.g. Fig. S3). In each iteration of the bootstrap, lineages were sampled with replacement from the full dataset. For a dataset comprising *n* lineages, *n* samples were taken. The autocorrelation of the sampled dataset was then calculated according to the procedure above. After 1000 iterations, the error associated with each lag value was measured by calculating the standard deviation of all correlation

corresponding to this lag. For presentation, the autocorrelation for the full dataset (i.e. without sampling) is plotted with error bars denoting $\pm$ 1 standard deviation around each point.

Independent Inference of Pulse Timing

We wished to verify that our observation of sub-exponential gene expression pulse timing in the WT *B. subtilis,* reduced *B. subtilis* and reconstituted *E. coli* systems was not due to a systematic error in our peak-calling pipeline (e.g. under-sampling of short events). We therefore identified a feature in the reporter expression autocorrelations that is present when pulses are timed, but not when they last for exponentially-distributed durations. Specifically, when gene expression pulses are timed, the reporter production rate is strongly *positive* at the onset of the pulse, and is highly likely—due to sub-exponential timing—a narrowly-distributed time later (i.e. on the trailing edge of the pulse). This tight succession of events leads to a pronounced dip below zero in the autocorrelation at a lag determined by the characteristic duration of pulses. This negative correlation is not evident if pulse durations are exponentially-distributed. In this case, pulse onset no longer determines when the pulse will end, making it so that the extreme values of the reporter derivative do not consistently 'line up' to generate the signature negative correlation.

Calculating the autocorrelations for the WT, reduced and reconstituted systems—under all conditions tested—reveals precisely this signature of timing (Fig. S20). Notably, this signature is absent in the corresponding constitutive segmentation reporters in the same cells and conditions (Fig. S20, black curves). We note that this evidence of timing does not depend in any way on calling discrete events. The marked sub-exponential timing of the gene expression pulses produced by stochastic competition leads to a characteristic and unusual pattern in the temporal correlations of the reporter expression that is readily detectable.

Measurement of Fraction ON

In the main text we present the fraction of cells in the SinR target-positive state as a function of IPTG concentration. For the first measurement, we simply define a threshold above the average intensity of NDL-423 or TMN-1159 cells in the absence of IPTG, and identify ON and OFF cells as being above or below this threshold, respectively.

We note, however, that this measurement may overestimate the sharpness of the transition from OFF to ON at the population level, as the IPTG-inducible promoters driving *sinI* expression may themselves respond to IPTG in a nonlinear fashion. To address this issue in the *E. coli* system, we measured the output of a P_{Lac}-GFP reporter construct at a range of IPTG concentrations. To this end, NDL-425 (Δ*motA glmS::*P_{RNA1}*-RFP phoA::*P_{Lac}*-GFP attP::Z2(lacIq tetR)-specR clpPX::FRT-Kan*) and NDL-423 were diluted from overnight LB cultures > $1{:}10^5$, grown at 37 C to $OD_{600} \leq 0.1$, spotted onto agarose pads and imaged for RFP and GFP expression. Several hundred cells were imaged for each condition. The resulting images were segmented using a modified MATLAB algorithm operating on the constitutive RFP signal, and GFP fluorescence intensities were extracted. The GFP intensity of NDL-425 at each IPTG concentration was used as a proxy for the SinI expression level in NDL-423. We then scattered these estimated SinI expression levels against the measured fraction of cells in the ON state in NDL-423 to generate the inset of Fig. 2R. This procedure revealed that, while correcting for the response characteristics

of the Lac promoter reduced the sharpness of the transition, the SinI-SinR system itself generates substantial ultrasensitivity.

Calculating of GFP production rates under saturating [IPTG]

In the mathematical derivations below (Mathematical Derivations section 9), we derive an upper bound on the variability of durations of SinI dominance over SinR that is inferred from reporter trace data. This calculation requires a measurement of the variability in GFP production rate for the maximally-induced $P_{synthR1}$-*gfp* reporter (i.e. the GFP production rate in NDL-423 under saturating IPTG conditions). To arrive at this measurement, we grew NDL-423 cells in the *E. coli* mother machine device in the presence of 500 uM or 1 mM IPTG. Image analysis and lineage tracking were performed as described above. Similar estimates of the variability in GFP production rate resulted from both experiments ($CV_{500\ uM} = 0.50$ and $CV_{1\ mM} = 0.59$).

The GFP production rate were defined as the total amount of GFP produced between frames, that is:

$$\frac{dGFP}{dt} = GFP_{Tot,i+1} - GFP_{Tot,i}$$

Where $GFP_{Tot,i}$ denotes the total amount of GFP in a cell in frame i. The values reported below were calculated from cells that (1) did not divide between frames i and $i+1$, and (2) had a cell area of less than 300 pixels ($\langle Area \rangle = 388$ pixels). This conditioning ensured that the results *underestimated* the actual variability in production rate, as some rate variability is explained by cell size. This, in turn, meant that our estimates of the variability in duration of SinI dominance were larger than the true value. We chose these assumptions in order to arrive at an upper bound: see the Mathematical Derivations portion of the Supplemental Materials for more discussion.

HaloTag experiments

3 mL of cells were grown to the beginning of exponential phase (OD_{600} 0.05 – 0.1) in LB medium. HaloTag TMR reagent (Promega) was then added to the medium, tubes were wrapped in foil to shield them from light, and cells were allowed to grow for a further 15 minutes to allow labeling to occur. Cells were then immediately applied to the top of a disposable 0.22 μm bottle filter, and 200 mL of fresh medium was sucked through the filter to wash away unreacted dye molecules. Cells were then washed off the surface of the filter, concentrated by centrifugation, and applied to an agar pad. Phase images, chaining reporter expression, and TMR signal were then measured in populations of 5,000 – 10,000 cells by automated tiling microscopy. Cells in the images were initially identified using a simple phase-based segmentation algorithm, and masks were refined through a combination of automatic cell clipping and manual curation. All cells at the periphery of images, where overlap between adjacent frames might lead to bleaching, were eliminated. Average fluorescent intensities within each cell were then calculated and correlated between the two fluorescent channels.

Calculation of average gene expression pulse profiles

In Fig. S13, we report the average gene expression pulse profiles for events occurring under different levels of IPTG induction. For a detailed description of how these traces were calculated, see[1]. In brief, we first perform a simple average of all traces in a given dataset (i.e. frame *j* of the average is the average of the *j*th frame of all observed peaks). Next, we register each leading edge of each peak to the leading edge of this 'reference' average. Finally, we perform another simple

averaging of the registered peaks to yield the final composite peak. We note that outlying peaks (i.e. peaks that deviated from the reference trace by more than 2.5 standard deviations at any point) were discarded in calculating this average. This practice does not appreciably change the shape of the average profiles, as the outlying peaks are by definition rare, however it does noticeably change the calculated variability around this average. As we do not make any arguments from the amount of variation around this average trace, we continue the practice of discarding outlying peaks.

Simulations in Figure 1 of the main text

The figures were generated by simulating the system described by the equations in Fig. 2A using an implementation of the Gillespie algorithm in MATLAB (MathWorks). The lifetimes of all components (τ values) were all set to 1, so that time units correspond to simulated cell generations (assuming dilution through growth is the main degradation pathway). The parameter *c* was set to 100, making complexing highly favored.

To create Fig. 1B, λ_R was set to 50 and λ_I was swept from 0 to 100 in steps of 2. At each value of λ_I, dynamics were simulated through 20,000 or more simulated generations, allowing deep sampling of the average number of free SinI and SinR molecules. The traces in Fig. 1B represent small sections of these traces with $\lambda_I = 30$, 42 or 50.

The ON Fraction in Fig. S2A was simulated similarly, and the ON state for production of *Z* was defined by the Heaviside function $H(R)$. The three traces show the behavior when λ_R is fixed at 50, 100, or 200, and λ_I is swept from 0 to 200% of λ_R.

The inter-peak times and peak durations shown in Figs. 1C and D were generated using essentially the same computational pipeline applied to simulated reporter traces of the system defined in Fig. 1B of the main text. However, because these traces offer perfect resolution, we purposefully remove any peaks called by peakfinder that are less than a generation apart (i.e. we replace two very close peaks with a single peak), as they would not be resolvable in the experimental systems. The autocorrelation functions shown in Fig. S3 were analytically derived for a process in which immature reporter is born in large bursts occurring with constant probability (rate λ_1) and matures with first-order kinetics (rate λ_2). Both mature and immature species decay exponentially with the same average lifetime (set to 1 cell generation to capture dilution). The maturation rate for immature protein was chosen to fit the derived curves to experimental data. We used average maturation times (i.e. reciprocal of exponential rate constant) of 5.6 min for GFPmut2 in the *E. coli* reconstitution circuit, 34 min for mKate2 in the reduced *B. subtilis* circuit and 6.6 min for CFP in the WT *B. subtilis* reporter strain. Values for fluorescent protein maturation rates were taken from a recent systematic survey of fluorescent protein maturation in *E. coli* (*28*). We note that maturation times for mKate2 and CFP have not been measured in *B. subtilis* to our knowledge. We therefore regard fits to the *B. subtilis* data as approximate.

In the derivation (see SM, Mathematical Derivations section 8), we derived the expected curve assuming large burst sizes so as to make random fluctuations in reporter maturation negligible. This effectively means that we follow only noise resulting from the coarse-grained stochastic competition between SinI and SinR. Analytical treatment and simulations with varying λ_1 reveal that the shape of the normalized autocorrelation function is invariant with respect to the burst

initiation rate. We therefore expect all data from the stochastic competition strains—irrespective of pulse firing frequency—to exhibit the same autocorrelation shape.

Stochastic antagonism between two proteins governs a bacterial cell fate switch (Materials and Methods — Mathematical Derivations)

In this supplemental document we provide the mathematical derivations underlying some of the statements made in the main text. The table of contents below can be used to find proofs of specific statements. However, we note that the document builds as it goes through, so some understanding of early sections may be necessary to understand later ones. To aid in interpretation, we open with an Outline that highlights the main results.

Contents

1 Outline

We will start with a quick overview to highlight the key properties of the system. Throughout the document, SinR is represented in equations by X and SinI by Y. Most of the mathematics below concerns attempts to quantify periods of dominance of SinI — i.e. the duration and variability in the amount of time that the SinI concentration exceeds that of SinR, so that SinR targets are derepressed.

- In Section 2, we describe how a stochastic model for the interaction of X and Y can be reduced to a single variable model describing the difference $Y - X$ when complexing occurs quickly (as in the native system)
- In Section 3, we use the simplified model to show that it impossible to achieve narrowly distributed ($CV < 1$) periods of SinI dominance without bursting in SinI, and that SinI dominance periods are generally very short (shorter than one cell division)
- In Section 4, we consider bursty production of both species. We first show a transformation that allows us to map a system with bursty production of both SinI and SinR onto a simplified model with bursting only in SinI. We then show how bursting is expected to affect the distribution and duration of the periods of SinI dominance.
- In Section 5, we show that initiation of SinI dominance periods is a memoryless process, and that the rate of initiation is highly sensitive to the the ratio of the production rates of the two proteins λ_Y/λ_X.
- The results of the previous sections show that the interaction between SinI and SinR is roughly governed by two parameters: (1) the ratio of the production rates of the two proteins λ_Y/λ_X and (2) the ratio of average burst sizes in production of the two proteins N_b/N_a. The first roughly sets the average frequency SinI dominance periods, and the second controls how narrowly distributed those are. However, the actual duration of those periods is fairly insensitive to both parameters, and they are typically very short. This means that the SinI-SinR circuit alone functions largely as a "pulse generator." We finally show in Section 6 that the addition of SlrR provides additional flexibility that can be used to tune the duration of the dominance periods.
- Section 7 explores a more mechanistically complex model that incorporates more fine-grained steps and reactions motivated by what is known about the biology and also the biochemical rate constants that have been measured. We show the properties of this more complicated model can be replicated using a matched reduced model, suggesting that the key properties and behaviors of the native circuit can be understood by considering the simpler models considered in the previous sections.
- Direct experimental observation of the mechanism is made difficult by the fact that switching occurs stochastically in single cells, is mediated by complexing of the two components (rather than, for example, presence or absence of a control molecule), and the fast timescale of the SinI-dominant state. In Sections 8 and 9 that examine how some of the expected properties can be inferred from experimental data.
- Finally, as some readers may find diffusion approximations to be more intuitive than the discrete point processes we focus on in most of this document, in Section 10 we briefly rederive some of the main results within this framework.

2 Modeling the SinI-SinR interaction

A simple model of constitutive production, exponential dilution/degradation, and complexing of SinR (below denoted X) and SinI (below denoted Y) follows the reactions

$$\begin{aligned} X &\xrightarrow{\lambda_X} X+1 & Y &\xrightarrow{\lambda_Y} Y+b \\ X &\xrightarrow{X/\tau_X} X-1 & Y &\xrightarrow{Y/\tau_Y} Y-1 \\ (X,Y) &\xrightarrow{cXY} (X-1,Y-1) \end{aligned} \tag{1}$$

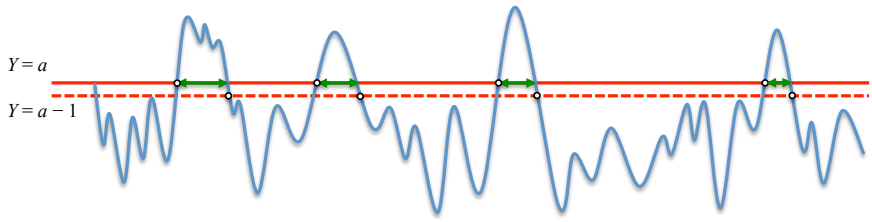


Figure 1: **First passage time calculation.** The figure shows the trajectory of Y during several consecutive switching events. Each period of Y dominance is defined by when Y first exceeds a particular threshold $a-1$ and then returns below it. The figure shows the case where Y is produced in steps of 1 ($b = 1$) so that the first jump above the threshold is always to a level of a molecules. Example dominance periods are shown with the green arrows. The duration of a particular dominance period is thus a random variable, which we denote by $T_{a\to a-1}$.

where λ_X and λ_Y are the production rates, τ_X and τ_Y are the half-lives, and c is the rate constant for complex formation. Production of Y is allowed to occur in bursts: b can potentially be a positive random variable [1]. We assume both proteins are subject to the same degradation pathway and/or dilution through cell growth, and thus have the same half-lives. For convenience we thus set $\tau_X = \tau_Y = 1$. For stable proteins, time units will then be in generations of growth. The relative magnitudes of λ_X and λ_Y will then set whichever species is in the majority on average.

As heterodimer formation is effectively irreversible, complexing becomes another elimination pathway for X and Y: as soon as free X and Y molecules are produced, they will immediately pair up, exhausting whichever species is less abundant. This creates states where one species dominates while the other one vanishes. Most of the time the majority species X will dominate because of its higher average abundance, but stochasticity in the reactions will occasionally lead to Y-dominant states. This spontaneous switching underlies the stochastic antagonism mechanism for phenotype switching.

In what follows, we will examine the durations of these temporary periods of dominance. In some parameter regimes, we will see that stochastic antagonism yields dominance periods that show greater-than-exponential variance, in contrast to expectations from, e.g., standard bistable switching mechanisms. We will also see that the addition of bursting in production can lead to markedly reduced variance in the dominance periods, meaning that stochastic antagonism can effectively impart timing without the need for feedback.

2.1 Simplified dynamics during a dominance period in the fast complexing limit

Assume that Y has temporarily become dominant. If complexing is efficient (c is very large), any X molecule that is produced immediately removes a free Y molecule: complexing effectively becomes instantaneous. In this limit, Y dynamics during the dominance period can be described by a reduced model:

$$Y \xrightarrow{\lambda_Y} Y + b \tag{2}$$

$$Y \xrightarrow{\lambda_X + Y} Y - 1 \tag{3}$$

This is possible since whenever $Y > 0$ in the fast complexing limit, $X = 0$: the model above thus can also be viewed as governing the dynamics of $Y - X$. A natural question to ask then is how long the period of Y dominance then persists. To do this, we will explicitly calculate the statistics of a first passage time: conditioned on starting at some point y_0, we ask how long the trajectory of Y remains above a barrier at the level $a-1$. The distribution of these times over subsequent periods of dominance will be denoted $T_{y_0\to a-1}$. Fig. (1) shows the definition of the involved quantities.

[1] Similar calculations can be done allowing both species to have bursty production, but we will here present the simpler results when X is produced in steps of 1.

In most cases it is safe to assume $a = 1$, so that we are asking how long it takes for the number of free molecules to reach 0. However, the results hold true for arbitrary $a > 1$ as well, which may be important, for example, when interpreting experimental data where other factors might effectively raise the threshold (e.g. if some SinI molecules are bound to other interacting proteins). Similarly, when there is no bursting in Y production ($b = 1$) we can take y_0 to be 1, meaning we are looking at when Y levels first exceed those of X. However, some calculations are done for arbitrary y_0, as we will generalize to the case where b is a random variable below.

3 In the absence of bursting, dominance periods of the minority species are short but show greater than exponential variation

We first focus on the non-bursting case $b = 1$. Assuming the barrier $a = 1$, the relevant first passage time is then $T_{1\to 0}$. In this regime, we will show that the dominance periods are quite noisy, showing greater than exponential variation. We do this by calculating the coefficient of variation of $T_{1\to 0}$, meaning that we must compute the mean $\langle T_{1\to 0}\rangle$ and variance $\mathrm{Var}(T_{1\to 0})$.

3.1 Deriving the backward equation

First passage times are typically computed using the backward master equation for the system, which we will here derive from first principles. We start from the Chapman-Kolmogorov equation [28, 29], with $P(y,t|y_0,t_0)$ denoting the probability of there being y molecules at time t under the initial condition of y_0 molecules at t_0:

$$\begin{aligned} P(y,t|y_0,t_0) =& \sum_{\hat{y}} P(y,t|\hat{y},t_0+\Delta t)P(\hat{y},t_0+\Delta t|y_0,t_0) \\ =& P(y,t|y_0+1,t_0+\Delta t)\lambda_Y\Delta t + P(y,t|y_0-1,t_0+\Delta t)(y_0+\lambda_X)\Delta t \\ &+ P(y,t|y_0,t_0+\Delta t)\left[1-\lambda_Y\Delta t-(y_0+\lambda_X)\Delta t\right]+o(\Delta t) \end{aligned}$$

where, under the usual definition for probability transition rate [28, 29], we have from (2) and (3) that $\lim_{\Delta t\to 0} P(y_0+1,t_0+\Delta t|y_0,t_0)/\Delta t \equiv \lambda_Y$ and $\lim_{\Delta t\to 0} P(y_0-1,t_0+\Delta t|y_0,t_0)/\Delta t \equiv y_0+\lambda_X$.

By moving the term $P(y,t|y_0,t_0+\Delta t)$ from right to left, dividing Δt on both sides, and letting Δt go to zero, we have

$$\begin{aligned} \lim_{\Delta t\to 0}\frac{P(y,t|y_0,t_0)-P(y,t|y_0,t_0+\Delta t)}{\Delta t} =& \lim_{\Delta t\to 0}\{\lambda_Y[P(y,t|y_0+1,t_0+\Delta t)-P(y,t|y_0,t_0+\Delta t)] \\ &+(y_0+\lambda_X)[P(y,t|y_0-1,t_0+\Delta t)-P(y,t|y_0,t_0+\Delta t)]+O(\Delta t)\} \\ \Rightarrow \quad -\partial_{t_0}P(y,t|y_0,t_0) =& \lambda_Y[P(y,t|y_0+1,t_0)-P(y,t|y_0,t_0)] \\ &+(y_0+\lambda_X)[P(y,t|y_0-1,t_0)-P(y,t|y_0,t_0)] \end{aligned} \tag{4}$$

Since the process is time-homogeneous (the rates are not time-dependent), this expression can be simplified by noting

$$P(y,t|y_0,t_0) = P(y,t-t_0|y_0,0) \quad . \tag{5}$$

Similarly, by changing variables for the partial derivative,

$$\partial_{t_0}P(y,t|y_0,t_0) = \partial_{t_0}P(y,t-t_0|y_0,0) = -\partial_t P(y,t|y_0,0) \quad . \tag{6}$$

We may also suppress the time of the initial condtion by defining $P(y,t|y_0) \equiv P(y,t|y_0,0)$, yielding a simplified version of the dynamics:

$$\begin{aligned} \partial_t P(y,t|y_0) =& \lambda_Y[P(y,t|y_0+1)-P(y,t|y_0)] \\ &+(y_0+\lambda_X)[P(y,t|y_0-1)-P(y,t|y_0)] \end{aligned} \tag{7}$$

3.2 Defining the barrier

As noted above, we here consider an absorbing boundary at $a-1$. The probability that the system starts at y_0 (at time 0) and remains in $[a,\infty)$ until time t is then given by the sum

$$G(y_0,t) \equiv \sum_{y=a}^{\infty} P(y,t|y_0,0) = \text{Prob}\{T_{y_0\to a-1} \geq t\} \tag{8}$$

where $T_{y_0\to a-1}$ denotes the time spend of Y molecules goes from y_0 to $a-1$. Here we will focus without loss of generality on the case $a=1$, corresponding to setting the threshold at zero Y molecules. The calculation for the general case is essentially identical.

By doing the same summation $\sum_{y_0=1}^{\infty}$ on both sides and using symbol y instead of y_0, Eq. (7) can be used to derive an evolution equation for $G(y_0,t)$:

$$\partial_t G(y_0,t) = \lambda_Y[G(y_0+1,t) - G(y_0,t)] + (y_0+\lambda_X)[G(y_0-1,t) - G(y_0,t)] \tag{9}$$

This relation can in turn be used to derive reccurence relations for the moments of the distributions of $T_{y_0\to 0}$, as we now describe.

3.3 A recurrence relation for the average first passage times $\left\langle T_{y_0\to 0}\right\rangle$

$G(y_0,t)$ is related to the first passage time through the relation $\text{Prob}\{T_{y_0\to 0}=t\} = -\partial_t G(y_0,t)$. The average can thus be calculated by integrating

$$\begin{aligned}
\left\langle T_{y_0\to 0}\right\rangle &= \int_{t=0}^{+\infty} tP\{T_{y_0\to 0}=t\}dt \\
&= \int_{t=0}^{+\infty} t[-\partial_t G(y_0,t)]dt \\
&= -tG(y_0,t)|_{t=0}^{t=+\infty} + \int_{t=0}^{+\infty} G(y_0,t)dt \\
&= \int_{t=0}^{+\infty} G(y_0,t)dt
\end{aligned} \tag{10}$$

We next integrate both sides of Eq. (9) in order to generate relations in terms of $\left\langle T_{y_0\to 0}\right\rangle$. The LHS is simply $\int_{t=0}^{+\infty}\partial_t G(y_0,t)dt = G(y_0,t)|_{t=0}^{t=+\infty} = 0-1$, so the full equation becomes (for arbitrary y_0):

$$\lambda_Y[\left\langle T_{y_0+1\to 0}\right\rangle - \left\langle T_{y_0\to 0}\right\rangle] + (y_0+\lambda_X)[\left\langle T_{y_0-1\to 0}\right\rangle - \left\langle T_{y_0\to 0}\right\rangle] = -1 \tag{11}$$

This result can be used to establish a recurrence relation for the first passage times with different possible starting points. To reduce notation, we define an auxiliary function

$$\phi(y_0) = \frac{y_0+\lambda_X}{\lambda_Y}\cdot\frac{(y_0-1)+\lambda_X}{\lambda_Y}\cdots\frac{1+\lambda_X}{\lambda_Y} \tag{12}$$

and $\phi(0)=1$. (11) can then be restated as

$$\lambda_Y\phi(y_0)\left[\frac{\left\langle T_{y_0+1\to 0}\right\rangle - \left\langle T_{y_0\to 0}\right\rangle}{\phi(y_0)} - \frac{\left\langle T_{y_0\to 0}\right\rangle - \left\langle T_{y_0-1\to 0}\right\rangle}{\phi(y_0-1)}\right] = -1 \tag{13}$$

The absorbing boundary condition can be written as $\langle T_{0\to 0}\rangle = 0$. We then have an infinite set of equations relating different starting heights (with $z > 0$ below an arbitrary height)

$$\begin{aligned}
\langle T_{0\to 0}\rangle &= 0 \\
&\cdots \\
\frac{\langle T_{z+1\to 0}\rangle - \langle T_{z\to 0}\rangle}{\phi(b)} - \frac{\langle T_{z\to 0}\rangle - \langle T_{z-1\to 0}\rangle}{\phi(b-1)} &= -\frac{1}{\lambda_Y \phi(b)} \\
\frac{\langle T_{z\to 0}\rangle - \langle T_{z-1\to 0}\rangle}{\phi(z-1)} - \frac{\langle T_{z-1\to 0}\rangle - \langle T_{z-2\to 0}\rangle}{\phi(z-2)} &= -\frac{1}{\lambda_Y \phi(z-1)} \\
&\cdots \\
\frac{\langle T_{3\to 0}\rangle - \langle T_{2\to 0}\rangle}{\phi(2)} - \frac{\langle T_{2\to 0}\rangle - \langle T_{1\to 0}\rangle}{\phi(1)} &= -\frac{1}{\lambda_Y \phi(2)} \\
\frac{\langle T_{2\to 0}\rangle - \langle T_{1\to 0}\rangle}{\phi(1)} - \frac{\langle T_{1\to 0}\rangle - \langle T_{0\to 0}\rangle}{\phi(0)} &= -\frac{1}{\lambda_Y \phi(1)}
\end{aligned} \tag{14}$$

3.4 A recurrence relation for the second moments $\left\langle T^2_{y_0\to 0}\right\rangle$

The calculation for the second moment is similar:

$$\begin{aligned}
\left\langle T^2_{y_0\to 0}\right\rangle &= \int_{t=0}^{+\infty} t^2[-\partial_t G(y_0,t)]dt = -t^2 G(y_0,t)|_{t=0}^{t=+\infty} + \int_{t=0}^{+\infty} 2tG(y_0,t)dt \\
&= \int_{t=0}^{+\infty} 2tG(y_0,t)dt
\end{aligned} \tag{15}$$

By multiplying $2t$ and then integrating t from 0 to ∞ on the left side of Eq. (9), we obtain (for y_0 inside $(0,\infty)$)

$$\int_{t=0}^{+\infty} 2t\partial_t G(y_0,t)dt = \int_{t=0}^{+\infty} G(y_0,t) = \langle T_{y_0\to 0}\rangle$$

and thus the relationship

$$\lambda_Y\left[\left\langle T^2_{y_0+1\to 0}\right\rangle - \left\langle T^2_{y_0\to 0}\right\rangle\right] + (y_0+\lambda_X)\left[\left\langle T^2_{y_0-1\to 0}\right\rangle - \left\langle T^2_{y_0\to 0}\right\rangle\right] = -2\langle T_{y_0\to 0}\rangle \tag{16}$$

Again the boundary condition for the second order moment follows $\langle T^2_{0\to 0}\rangle = 0$. Following the same procedure with an identical definition of ϕ

$$\begin{aligned}
\left\langle T^2_{0\to 0}\right\rangle &= 0 \\
&\cdots \\
\frac{\langle T^2_{z+1\to 0}\rangle - \langle T^2_{z\to 0}\rangle}{\phi(b)} - \frac{\langle T^2_{z\to 0}\rangle - \langle T^2_{z-1\to 0}\rangle}{\phi(b-1)} &= -\frac{2\langle T_{z\to 0}\rangle}{\lambda_Y \phi(b)} \\
\frac{\langle T^2_{z\to 0}\rangle - \langle T^2_{z-1\to 0}\rangle}{\phi(z-1)} - \frac{\langle T^2_{z-1\to 0}\rangle - \langle T^2_{z-2\to 0}\rangle}{\phi(z-2)} &= -\frac{2\langle T_{z-1\to 0}\rangle}{\lambda_Y \phi(z-1)} \\
&\cdots \\
\frac{\langle T^2_{3\to 0}\rangle - \langle T^2_{2\to 0}\rangle}{\phi(2)} - \frac{\langle T^2_{2\to 0}\rangle - \langle T^2_{1\to 0}\rangle}{\phi(1)} &= -\frac{2\langle T_{2\to 0}\rangle}{\lambda_Y \phi(2)} \\
\frac{\langle T^2_{2\to 0}\rangle - \langle T^2_{1\to 0}\rangle}{\phi(1)} - \frac{\langle T^2_{1\to 0}\rangle - \langle T^2_{0\to 0}\rangle}{\phi(0)} &= -\frac{2\langle T_{1\to 0}\rangle}{\lambda_Y \phi(1)}
\end{aligned} \tag{17}$$

3.5 An exact inequality for the non-bursting case $b = 1$

The systems of equations given by 14 and 17 can be solved in the general case (arbitrary y_0), but the answers are quite involved. Here we focus on the non-bursting case ($b = 1$ in 2). In this case the initial burst will always push the system only one unit over the barrier, so the relevant statistics are $\langle T_{1\to 0}\rangle$ and $\langle T_{1\to 0}^2\rangle$.

$\langle T_{1\to 0}\rangle$ can be calculated by noting that the sum of the equations in 14 is essentially a telescoping sum showing that

$$\begin{aligned}\langle T_{1\to 0}\rangle &= \sum_{y_0=1}^{\infty} \frac{1}{\lambda_Y \phi(y_0)} && (18)\\ &= \frac{1}{\lambda_Y}\sum_{y_0=1}^{\infty} \frac{\lambda_Y^{y_0}}{[y_0+\lambda_X][(y_0-1)+\lambda_X]\cdots[1+\lambda_X]} \\ &= \frac{1}{\lambda_Y}\lambda_X e^{\lambda_Y}\lambda_Y^{-\lambda_X}\gamma(\lambda_X,\lambda_Y) - \frac{1}{\lambda_Y} && (19)\end{aligned}$$

where $\gamma(s, y_0) \equiv \int_0^{y_0} e^{-t}t^{s-1}dt$ is the lower incomplete gamma function.

Likewise, we can obtain a summed expression for $\langle T_{1\to 0}^2\rangle$:

$$\left\langle T_{1\to 0}^2\right\rangle = \sum_{y_0=1}^{\infty} \frac{2\left\langle T_{y_0\to 0}\right\rangle}{\lambda_Y \phi(y_0)}$$

To simplify this we observe that each of the steps $T_{x\to x-1}$ (the first passage time from x to $x-1$) occurs independently

$$\begin{aligned}\left\langle T_{y_0\to 0}\right\rangle &= \sum_{x=1}^{y_0}\langle T_{x\to x-1}\rangle = \langle T_{1\to 0}\rangle + \sum_{x=2}^{y_0}\langle T_{x\to x-1}\rangle \\ &= \langle T_{1\to 0}\rangle + \left\langle T_{y_0\to 1}\right\rangle && (20)\end{aligned}$$

This relationship can be used to bound the moment:

$$\begin{aligned}\left\langle T_{1\to 0}^2\right\rangle &= \sum_{y_0=1}^{\infty} \frac{2\left\langle T_{y_0\to 0}\right\rangle}{\lambda_Y \phi(y_0)} \\ &= 2\langle T_{1\to 0}\rangle \sum_{y_0=1}^{\infty} \frac{1}{\lambda_Y \phi(y_0)} + \sum_{y_0=2}^{\infty} \frac{2\left\langle T_{y_0\to 1}\right\rangle}{\lambda_Y \phi(y_0)} && (21)\\ &= 2\langle T_{1\to 0}\rangle^2 + \sum_{y_0=2}^{\infty} \frac{2\left\langle T_{y_0\to 1}\right\rangle}{\lambda_Y \phi(y_0)} && (22)\\ &> 2\langle T_{1\to 0}\rangle^2 && (23)\end{aligned}$$

where from (21) to (22), we use Eq. (18).

Therefore the coefficient of variance (CV) of the dominance period duration for the non-bursting process obeys

$$\mathrm{CV} = \frac{\sqrt{\mathrm{Var}[T_{1\to 0}]}}{\langle T_{1\to 0}\rangle} = \frac{\sqrt{\left\langle T_{1\to 0}^2\right\rangle - \langle T_{1\to 0}\rangle^2}}{\langle T_{1\to 0}\rangle} > \frac{\sqrt{2\langle T_{1\to 0}\rangle^2 - \langle T_{1\to 0}\rangle^2}}{\langle T_{1\to 0}\rangle} = 1 \quad (24)$$

This result means that dominance periods in the non-bursting model cannot be narrower than the exponential distribution, which has a CV of 1.

3.6 Numerical extension to the slower complexing regime

The exact inequality above for the fast complexing limit in fact is a lower bound on overall behavior. Numerical simulations show that as c becomes smaller (so that dilution becomes more important), the CV is even larger than when complexing dominates (see Fig. 2). Examination of the moments shows that this behavior occurs because the mean residence time decreases monotonically with c, while the second moment remains almost the same.

3.7 In the absence of bursting, dominance periods of the minority species are short

The moment equations Eq. (18) also allow us to bound the average duration that SinI is dominant. Again as in Fig. 1, we define the dominance period as the duration over which Y exceeds the barrier at $a-1$. If there is no bursting, then the beginning of the dominance period always starts with a molecules, and on average lasts

$$\begin{aligned}
\langle T_{a\to a-1}\rangle &= \frac{1}{a+\lambda_X} + \frac{\lambda_Y}{a+\lambda_X}\cdot\frac{1}{(a+1)+\lambda_X} + \frac{\lambda_Y}{a+\lambda_X}\cdot\frac{\lambda_Y}{(a+1)+\lambda_X}\cdot\frac{1}{(a+2)+\lambda_X} + \cdots \\
&< \frac{1}{a+\lambda_X} + \frac{\lambda_Y}{a+\lambda_X}\frac{1}{a+\lambda_X} + \left(\frac{\lambda_Y}{a+\lambda_X}\right)^2\frac{1}{a+\lambda_X} + \cdots \\
&= \frac{1}{a+\lambda_X}\left[1 + \frac{\lambda_Y}{a+\lambda_X} + \left(\frac{\lambda_Y}{a+\lambda_X}\right)^2 + \left(\frac{\lambda_Y}{a+\lambda_X}\right)^3 + \cdots\right] \\
&= \frac{1}{a+\lambda_X}\frac{1}{1-\frac{\lambda_Y}{a+\lambda_X}} = \frac{1}{a+\lambda_X}\frac{a+\lambda_X}{a+\lambda_X-\lambda_Y} \\
&= \frac{1}{a+\lambda_X-\lambda_Y} \quad . \qquad (25)
\end{aligned}$$

Because we have rescaled time, time units are in generations. Since $a \geq 1$ and $\lambda_X \geq \lambda_Y$ by assumption, the expression above is thus bounded above by 1, so the dominance periods of the minority species are at most one generation long. Depending on the precise values of these parameters, dominance periods may in fact be much shorter on average.

Eq. 25 also serves as an estimate for the duration of SinI dominance periods. As seen in Fig. 3, the bound performs well and dominance periods are typically substantially shorter than a generation. Note there is still a strong dependence of the duration on the ratio of production rates λ_Y/λ_X.

4 Bursty production of the minority species can lead to short, narrowly distributed dominance periods

We now consider the effect of bursts in the production of the minority species. We will again consider the fast complexing limit. In this case, we allow b in Eq. 2 to be arbitrarily distributed. This is equivalent to considering a model where Y can be produced in many distinct production reactions and a single elimination reaction:

$$\begin{aligned}
Y &\xrightarrow{P(b)\lambda_Y} Y + b \quad \forall b > 0 \\
Y &\xrightarrow{\lambda_X + Y} Y - 1
\end{aligned}$$

where $P(b)$ denotes the probability that b molecules are produced in a given burst. We first show that this model is a actually a suitable simplification of a more complex model in which both species are produced in bursts, allowing us to proceed without loss of generality.

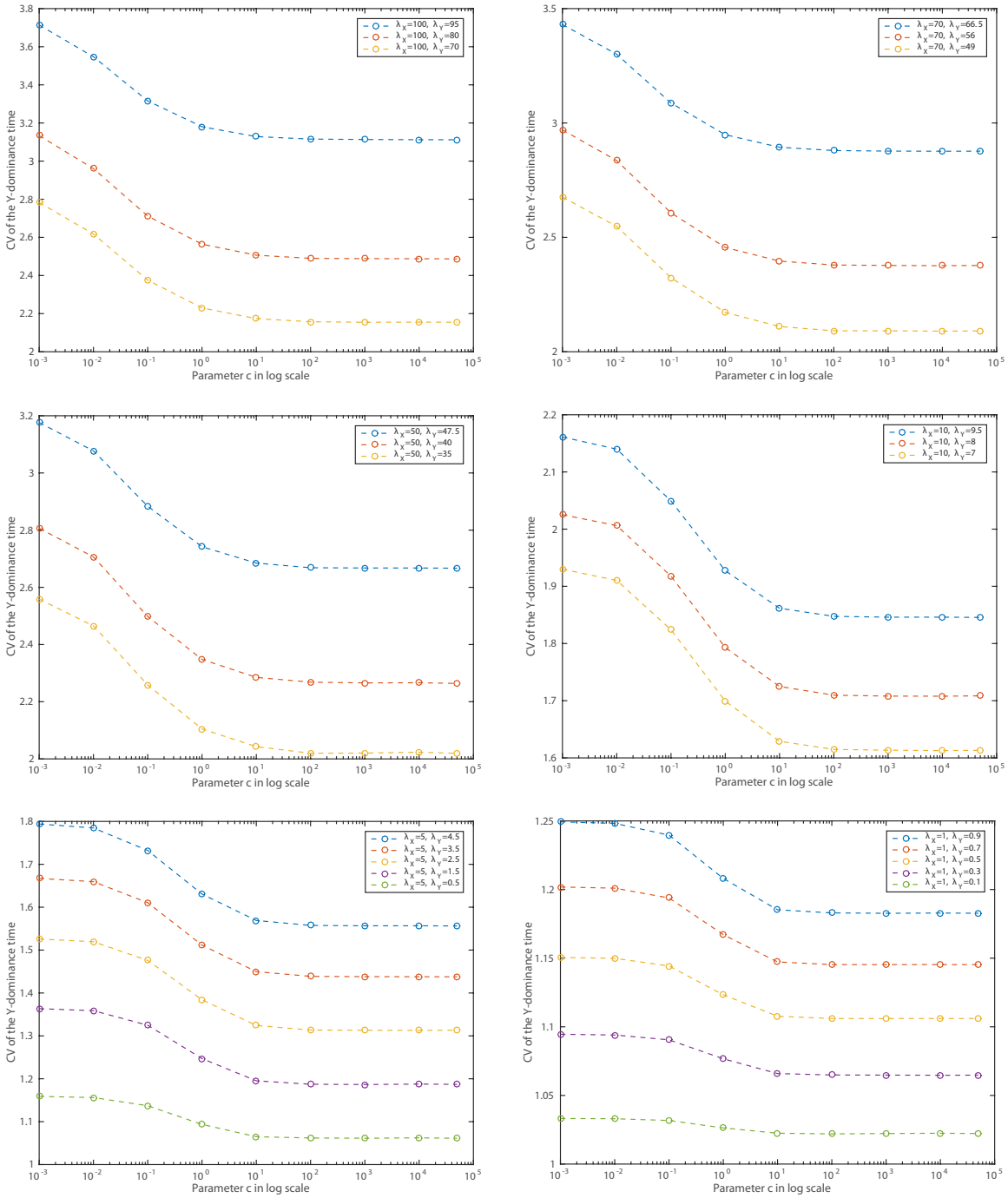


Figure 2: **Dependence of variability in Y-dominance on strength of complexing.** In the fast complexing limit, Eq. 24 shows that if the minority component Y is not produced in bursts, then the periods of Y-dominance will always be highly variable (have a $CV > 1$). The figure shows that this result is in fact a conservative bound on the true behavior: when complexing is not infinitely fast, the observed coefficients of variation are in fact higher than in the limiting case. Each panel plots the observed CV as c is changed for a a series of different pairs of production rates λ_X, λ_Y.

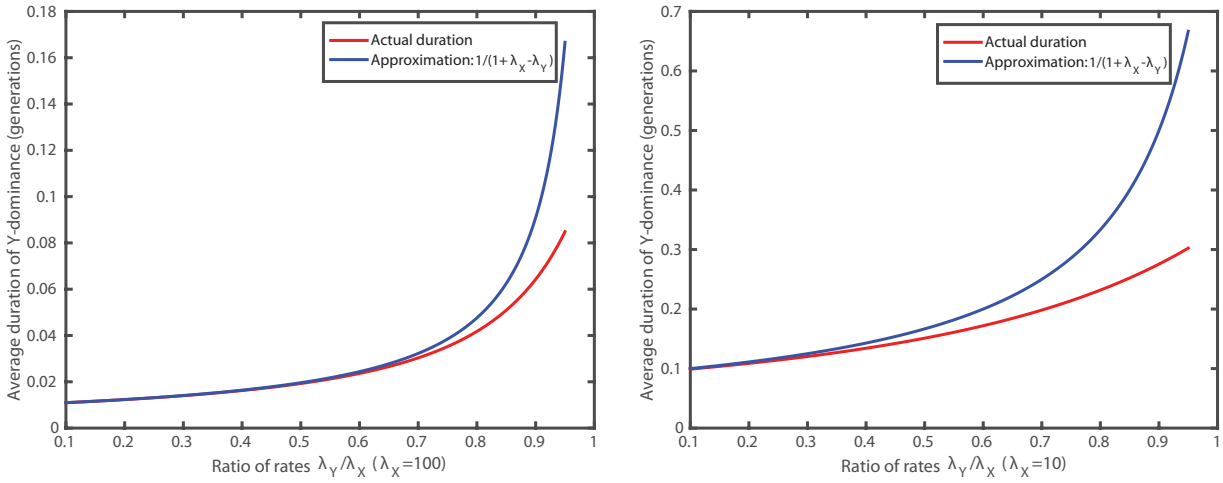


Figure 3: **Approximation for SinI dominance period durations.** In the fast complexing limit, Eq. 25 yields an estimate of the average duration of dominance of the minority species Y. The figure compares the bound to the actual dominance periods in two regimes as the ratio of production rates λ_Y/λ_X is changed.

4.1 A reduced model for bursting

As before, we denote SinR by X and SinI by Y. Suppose each mRNA molecule of X (or Y) is produced with rate λ_X (or λ_Y). The lifetime of each mRNA molecule is determined by the first binding event of a ribonuclease, leading to degradation. This lifetime is generally much smaller than the cell division time and can be considered to be instantaneous. The number of proteins generated by a single mRNA molecule is a result of individual ribosome binding events that competes with the binding of the degradation machinery. Denoting ρ as the probability of ribosome binding and $1-\rho$ for RNase E binding, the number of X proteins a produced from a single mRNA then follows a geometric distribution:

$$P_{N_a}(a) = \rho^a(1-\rho),\ a = 0,1,2,\ldots \tag{26}$$

where the parameter $N_a = \rho/(1-\rho)$ is the mean value of the distribution. A "contracted" version of SinI/SinR protein dynamics with bursting in both species then follows:

$$\begin{aligned} X &\xrightarrow{P_{N_a}(a)\cdot\lambda_X} X + a,\ \forall a \geq 0 \qquad Y \xrightarrow{P_{N_b}(b)\cdot\lambda_Y} Y + b,\ \forall b \geq 0 \\ X &\xrightarrow{X/\tau_X} X - 1 \qquad Y \xrightarrow{Y/\tau_Y} Y - 1 \\ (X,Y) &\xrightarrow{cXY} (X-1, Y-1) \end{aligned} \tag{27}$$

where λ_X and λ_Y are the production rates of mRNA molecules, τ_X and τ_Y are the half-lives of X and Y proteins, and c is the rate constant for complex formation. Note that there are actually an infinite number of production reactions above, as production of X and Y proteins occurs in geometrically-distributed bursts (the random variables a and b) with average burst sizes N_a and N_b. We assume both proteins are subject to the same degradation pathway and/or dilution through cell growth, and thus have the same half-lives. For convenience we thus set $\tau_X = \tau_Y = 1$.

Though the model has an infinite number of reactions, it can still be simulated efficiently through a modified Gillespie algorithm using mixture distributions. Through a large amount of simulation results in the region $\lambda_Y \leq \lambda_X$ (primarily the rare bursting regime) and $N_b \geq N_a$ (i.e. SinI is more bursty than SinR), we observed that the two parameters N_a and N_b can be reduced to be 1 and N_b/N_a with negligible influence on the residence time distribution of the Y-dominant state, as seen in Figure 4.

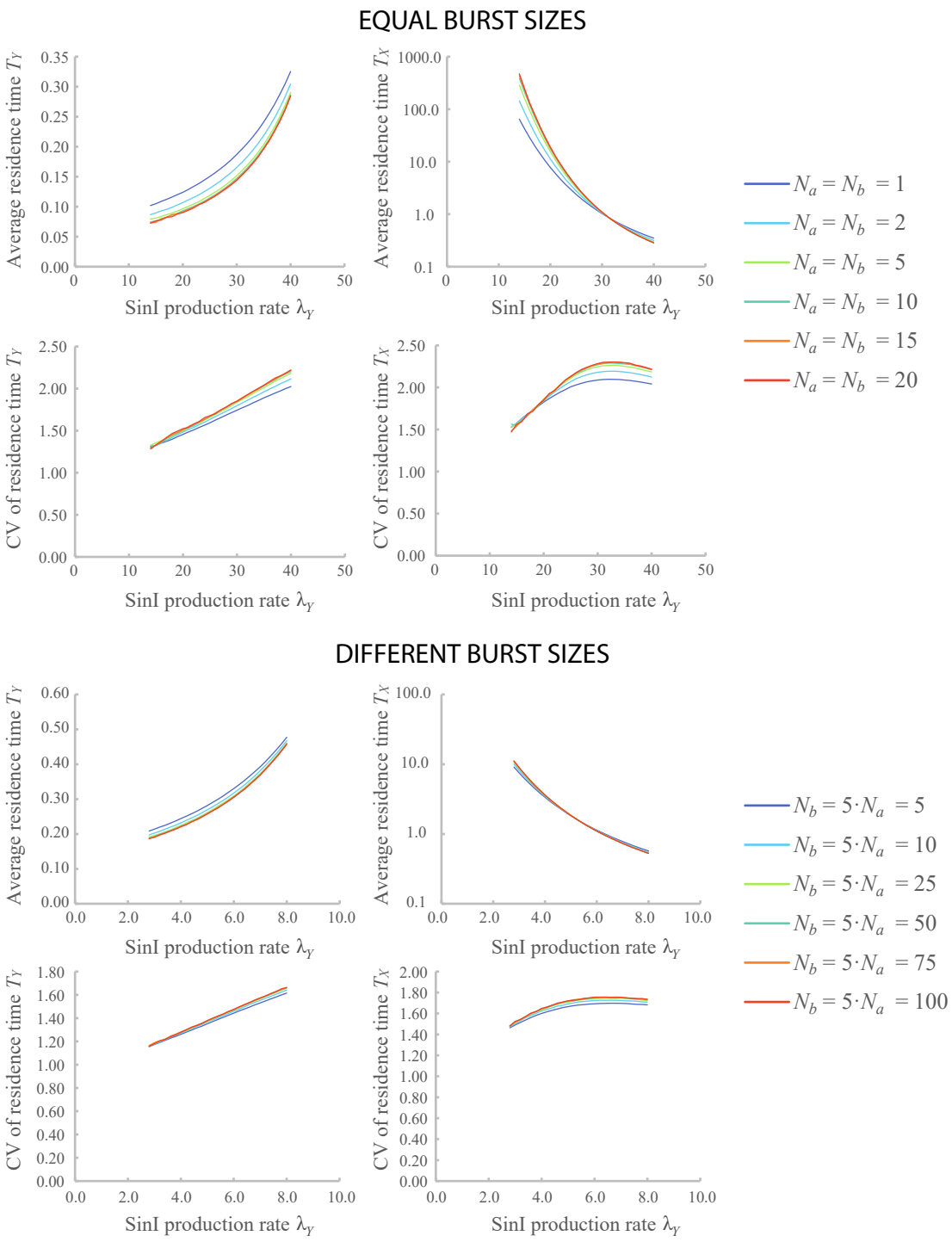


Figure 4: Representative simulations showing that it is primarily the ratio N_b/N_a that determines the statistical properties of pulses. The time spent in the$Y$-dominant (respectively X-dominant) state is referred to as T_Y (T_X). The simulations fix the production rate of X at 40, and vary the production rate of Y, which is denoted λ_Y. The first four panels summarize simulations in which both X and Y are produced in bursts with the same burst size. The first row of panels show the average time spent in the Y-dominant state (left) and the X-dominant state (right). The second row of panels shows the coefficient of variation for these residence times. In all cases the results are similar even though the burst size increases from 1 to 20. The second set of four panels shows similar simulations, except that the burst size in Y production has been set to be five times that in X production. Across a 20-fold change in burst size, the statistics are basically unchanged.

The model Eq. (27) can thus be reduced to one in which Y is produced in bursts and X is produced randomly one at a time:

$$\begin{aligned} X &\xrightarrow{\lambda_X} X+1 \qquad & Y &\xrightarrow{P_{N_b/N_a}(b)\cdot\lambda_Y} Y+b,\ \forall b \geq 0 \\ X &\xrightarrow{X} X-1 \qquad & Y &\xrightarrow{Y} Y-1 \\ (X,Y) &\xrightarrow{cXY} (X-1,Y-1) \end{aligned} \tag{28}$$

In the fast complexing limit, this model is in turn well-approximated by a model similar to the one discussed in Section 2.1:

$$\begin{aligned} Y &\xrightarrow{P_{N_b/N_a}(b)\cdot\lambda_Y} Y+b,\ \forall b \geq 0 \\ Y &\xrightarrow{\lambda_X+Y} Y-1 \end{aligned} \tag{29}$$

In the rare bursting limit, the the mean and variance of the Y-dominance time can then be calculated as mixture distributions (with $N' := N_b/N_a$):

$$\begin{aligned} \langle T_Y \rangle &= \sum_{b'=0}^{\infty} P_{N'}(b') \langle T_{b'+1\to 0} \rangle \\ &= \sum_{b'=0}^{\infty} \rho^{b'}(1-\rho) \sum_{y=1}^{b'+1} \frac{1}{\lambda_X + y} \\ &= \sum_{b'=0}^{\infty} \left(\frac{N'}{1+N'}\right)^{b'} \frac{1}{1+N'} \sum_{y=1}^{b'+1} \frac{1}{\lambda_X + y} \end{aligned} \tag{30}$$

$$\begin{aligned} \mathrm{Var}(T_Y) &= \sum_{b'=0}^{\infty} P_{N'}(b') \cdot \left\{ (\langle T_{b'+1\to 0} \rangle - \langle T_Y \rangle)^2 + \mathrm{Var}[T_{b'+1\to 0}] \right\} \\ &= \sum_{b'=0}^{\infty} \frac{N'^{b'}}{(1+N')^{b'+1}} \cdot \left\{ \left(\sum_{y=1}^{b'+1} \frac{1}{\lambda_X + y} - \langle T_Y \rangle \right)^2 + \sum_{y=1}^{b'+1} \left(\frac{1}{\lambda_X + y} \right)^2 \right\} \end{aligned} \tag{31}$$

An example comparison between simulations and the theoretical predictions for the reduced model is shown in Figure 5.

4.2 Bursts of a fixed size in the rare bursting limit

To make progress, we further specialize to the case where $\lambda_X \gg \lambda_Y$. This "rare bursting" limit ensures that each dominance period of Y corresponds to a unique burst in Y production that subsequently dilutes out. For example, say b molecules of Y have been produced. They will then each have exponentially distributed lifetimes: the first molecule will take $\mathrm{Exp}(\lambda_X + b)$ units of time to eliminate (where $\mathrm{Exp}(\lambda)$ denotes an exponential random variable with rate λ), the second molecule $\mathrm{Exp}(\lambda_X + b - 1)$ units of time, etc. The total time for all molecules to be eliminated will then follow the hypoexponential (generalized Erlang) distribution: $T_{b\to 0} \sim \mathrm{Hypo}(\lambda_X + b, \lambda_X + b - 1, \ldots, \lambda_X + 1)$. The mean and variance of this distribution follow:

$$\langle T_{b\to 0} \rangle = \sum_{y=1}^{b} \frac{1}{\lambda_X + y} \tag{32}$$

$$\mathrm{Var}(T_{b\to 0}) = \sum_{y=1}^{b} \left(\frac{1}{\lambda_X + y} \right)^2 \tag{33}$$

By Jensen's inequality we then have $\sum_{y=1}^{b} \left(\frac{1}{\lambda_X+y}\right)^2 \leq \left(\sum_{y=1}^{b} \frac{1}{\lambda_X+y}\right)^2$, implying $CV(T_{b\to 0}) \leq 1$ always, i.e. bursts are more narrowly distributed than the exponential expectation.

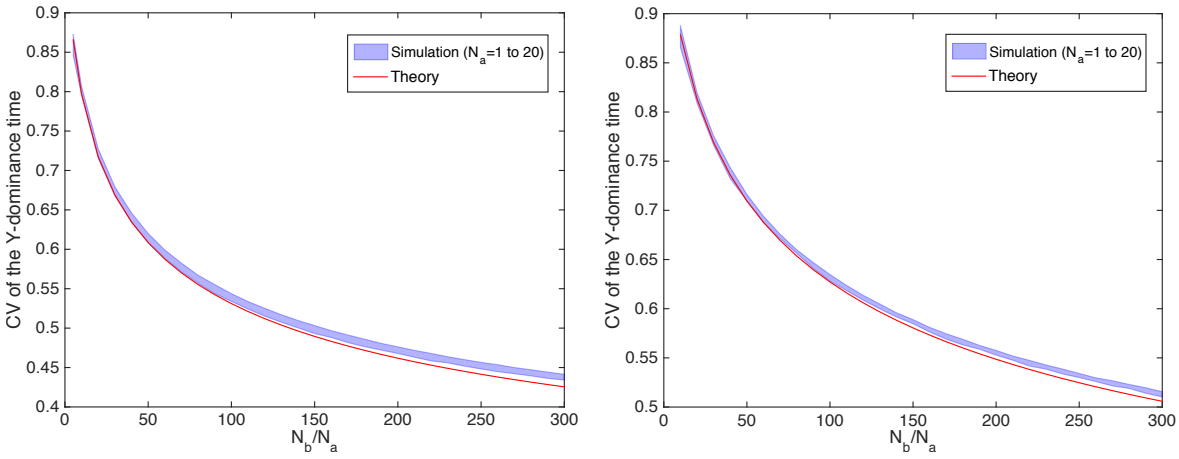


Figure 5: The simulations were performed by scanning N_a from 1 to 20 while keeping N_b/N_a as a fixed ratio. The theoretical curves are calculated using Eqs. (30) and (31). In the left panel $\lambda_Y = 0.01$ and $\lambda_X = 20$. In the left panel $\lambda_Y = 0.01$ and $\lambda_X = 50$.

4.3 Extension to distributed bursts

Most models of stochastic gene expression do not predict bursts of fixed size, but rather geometrically distributed bursts of protein production as individual mRNAs are produced and destroyed as discussed above. b is therefore a random variable. As the size of a particular burst may now vary, we simply denote the dominance duration by T_{burst}. The moments in this case can then be computed using the formulas for the moments of a mixture distribution:

$$\langle T_{\text{burst}} \rangle = \sum_{b'=0}^{\infty} P(b') \langle T_{b'+1\to 0} \rangle \ , \tag{34}$$

$$\text{Var}(T_{\text{burst}}) = \sum_{b'=0}^{\infty} P(b') \cdot \left\{ (\langle T_{b'+1\to 0} \rangle - \langle T_{burst} \rangle)^2 + \text{Var}(T_{b'+1\to 0}) \right\} \ . \tag{35}$$

Here b' is the "net burst". When the burst initiating a period of Y dominance occurs it produces b molecules, but any free X molecules must first be eliminated, leaving behind b' molecules. Connecting this distribution to that of b is generally impossible. However, in the regime we consider, production events of Y are rare, so we can regard X as being at steady state and in the dominant position (so that no free Y molecules are present). As the burst occurs, the average "net burst" is then $b' = b - \langle X \rangle$. In the case of the geometric distribution, we can then exploit memorylessness to see that

$$\begin{aligned} P(b') = P(b|b - \langle X \rangle \geq 0) &= \frac{P(b)}{P(b - \langle X \rangle \geq 0)} = \frac{(1-\rho)\rho^b}{1 - \sum_{y=1}^{\langle X \rangle - 1}(1-\rho)\rho^y} = \frac{(1-\rho)\rho^b}{\rho^{\langle X \rangle}} \\ &= (1-\rho)\rho^{b-\langle X \rangle} = (1-\rho)\rho^{b'} \end{aligned} \tag{36}$$

so the "net burst" has the same distribution as b. We will thus omit the prime in the subsequent discussions.

4.4 Two routes to narrowly distributed dominance periods

Bursting allows for accurate timing of the dominance periods in two different ways.

First, consider the case where $\lambda_X \gg \langle b \rangle$—what might be called the "small burst" regime. The exact moments can be calculated using 34 and 35, but the expressions are not particularly illuminating. We thus offer a more intuitive picture. In this regime complexing is the dominant elimination pathway. The lifetime of each Y molecule is then well-approximated as $\text{Exp}(\lambda_X)$. If there are b total molecules in a particular

burst, the total time to eliminate all molecules is then simply $T_{b\to 0} \sim \text{Gamma}(b, \lambda_X)$, which will be tightly concentrated around $\langle T_{b\to 0}\rangle = b/\lambda_X$. (Indeed, the gamma becomes a better and better approximation for the hypoexponential distribution discussed in Sec. 4.2 as λ_X increases.) If b is instead randomly distributed on $\{1, 2, 3, \cdots\}$ according to the distribution $P(b)$, $T_{b\to 0}$ will thus be a superposition of gamma distributions centered on the rescaled set $\{\frac{1}{\lambda_X}, \frac{2}{\lambda_X}, \frac{3}{\lambda_X}, \cdots\}$, and thus obey roughly the same distribution defined on $(0, \infty)$. As the CV is invariant under scaling, a low CV burst distribution will thus yield a low CV distribution for the duration of dominance periods $T_{b\to 0}$.

The second approach is to go to the opposite extreme and have extremely large bursts ($b \gg \lambda_X$). Dilution is then the dominant elimination pathway, and the same exponential dilution principle discussed in the main text takes over. Indeed, downstream targets of the SinI-SinR interaction can potentially benefit from two rounds of this effect: one at the level of bursting of SinI, and a second at the level of production of the targeted protein.

4.5 Durations of dominance periods when the minority species is produced in bursts

In the small burst regime, the approximate logic above suggests that the duration of dominance of the minority species is approximately

$$\sum_b P(b)\,\langle T_{b\to 0}\rangle \approx \sum_b P(b)(b/\lambda_X) = \langle b\rangle\,/\lambda_X \ll 1 \quad ,$$

so dominance periods are again shorter than a generation. 34 can be used to provide a more precise upper bound when the burst distribution is geometric:

$$\begin{aligned}
\langle T_{Burst}\rangle &= \sum_{b=0}^{\infty} P(b)\,\langle T_{b+1\to 0}\rangle = \sum_{b=0}^{\infty}(1-\rho)\rho^b \sum_{x=1}^{b+1}\frac{1}{\lambda_X + x} \\
&= \sum_{x=1}^{\infty}\frac{1}{\lambda_X + x}\sum_{b=x-1}^{\infty}(1-\rho)\rho^b = \sum_{x=1}^{\infty}\frac{1}{\lambda_X + x}\rho^{x-1} \\
&< \frac{1}{\lambda_X + 1}\sum_{x=0}^{\infty}\rho^x = \frac{1}{\lambda_X + 1}\frac{1}{1-\rho} = \frac{1}{\lambda_X + 1}\frac{1}{1 - \langle b\rangle/(\langle b\rangle + 1)} = \frac{\langle b\rangle + 1}{\lambda_X + 1}
\end{aligned} \tag{37}$$

5 Switches into the minority-dominant state occur memorylessly when rare

When $\lambda_X \gg \lambda_Y$, the majority species will dominate most of the time and switches into the minority-dominant state will be extremely rare. As they are also extremely short (see above), the likelihood of two overlapping bursts becomes smaller and smaller, and they are effectively impulsive events compared to the periods of in which the majority species is dominance. It is thus intuitively clear that switches into the minority state arise memorylessly, and occur according to a Poisson process. Here we will demonstrate that this is the case analytically by showing that the moments of the majority dominance periods tend to those of the exponential distribution. Since experimentally we can only measure when minority dominance periods begin (as the reporter responds slowly), we then show that this exponential distribution can be measured from the distribution of interpulse times (i.e. the distribution of waiting times from the start of one pulse to the start of the next).

Note that deviation from exponential interevent statistics is expected as λ_X and λ_Y become comparable, as in this regime it takes longer for a buffer of free molecules of either component to accumulate. This can lead to the system being primed to stay in one state, and thus coupling (and loss of memorylessness) between subsequent switching periods.

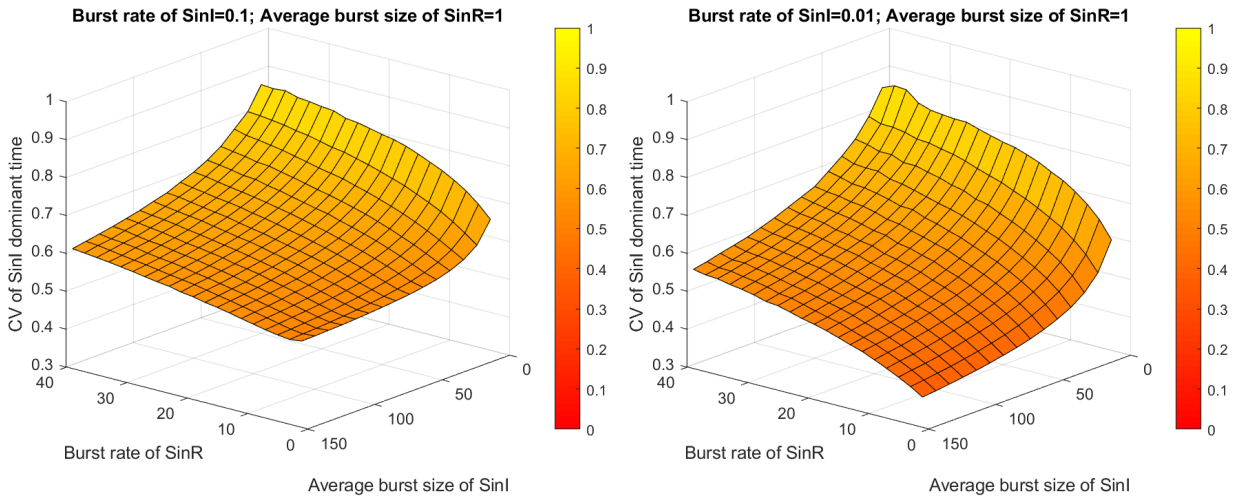


Figure 6: **Phase diagram for both SinI and SinR proteins produced in bursts.** There are four parameters in this model: the production rates of bursts for SinI and SinR and average burst sizes of SinI and SinR. But effectively the number of parameters is two, since the production rate of bursts for SinI is small and only the ratio of the average burst sizes between SinI and SinR influences the dynamics. We thus vary the average burst size of SinI and burst rate of SinR in the x-y plane and the in the z-axis show the CV for SinI dominance duration.

5.1 Periods of dominance for the majority species are exponentially distributed when $\lambda_X \gg \lambda_Y$

We will assume λ_Y is small, say $\lambda = \epsilon$ where $0 < \epsilon \ll \lambda_X$. As the system is symmetrical with respect to the two variables, a reduced model for the dynamics of X similar to Eqs. 2 and 3 can be derived, leading to an equation similar to Eq. 9. Again, this reduction to a single variable model is possible because fast complexing guarantees that when $X > 0$, we must have $Y = 0$. The simplified model thus in fact describes the dynamics of the difference $X - Y$. Switching will occur when this difference becomes negative, i.e. Y starts to dominate. We are thus interested in the first passage time $T_{x_0 \to -1}$ (i.e. the time to go from a surplus of x_0 SinR molecules to a deficit of one, when a period of Y dominance begins).

Similar to the derivation of Eq. 11 and Eq. 16, we can then develop a series of relationships among the moments of the first passage time $T_{x_0 \to -1}$

$$\lambda_X[\langle T_{x_0+1\to-1}\rangle - \langle T_{x_0\to-1}\rangle] + (x_0+\varepsilon)[\langle T_{x_0-1\to-1}\rangle - \langle T_{x_0\to-1}\rangle] = -1\,, \tag{38}$$

$$\lambda_X\left[\left\langle T^2_{x_0+1\to-1}\right\rangle - \left\langle T^2_{x_0\to-1}\right\rangle\right] + (x_0+\varepsilon)\left[\left\langle T^2_{x_0-1\to-1}\right\rangle - \left\langle T^2_{x_0\to-1}\right\rangle\right] = -2\left\langle T_{x_0\to-1}\right\rangle\,, \tag{39}$$

$$\cdots$$

$$\lambda_X\left[\left\langle T^n_{x_0+1\to-1}\right\rangle - \left\langle T^n_{x_0\to-1}\right\rangle\right] + (x_0+\varepsilon)\left[\left\langle T^n_{x_0-1\to-1}\right\rangle - \left\langle T^n_{x_0\to-1}\right\rangle\right] = -n\left\langle T^{n-1}_{x_0\to-1}\right\rangle\,, \tag{40}$$

$$\cdots$$

We will solve these equations for generic starting point x_0, but we are interested specifically in the first passage time $T_{0\to-1}$, meaning the time from when the two species are perfectly balanced (surplus of 0 SinR molecules) to when SinI first begins to dominate (a surplus of one SinI molecule, or equivalently a deficit of -1 SinR molecules). As before, we proceed by introducing an auxiliary function ϕ that allows the solutions to be phrased more simply. We set

$$\phi(x_0) = \frac{x_0+\varepsilon}{\lambda_X}\cdot\frac{(y_0-1)+\varepsilon}{\lambda_X}\cdots\frac{1+\varepsilon}{\lambda_X}\cdot\frac{\varepsilon}{\lambda_X} \tag{41}$$

and $\phi(-1) = 1$. From (38) we derive

$$\begin{aligned}\left\langle T_{x_0\to -1}\right\rangle &= \phi(x_0-1)\sum_{x=x_0}^{\infty}\frac{1}{\lambda_X\phi(x)} &(42)\\ &= \sum_{x=x_0}^{\infty}\frac{\lambda_X^{x-x_0}}{[x+\varepsilon][(x-1)+\varepsilon]\cdots[x_0+\varepsilon]}\\ &= \begin{cases} \frac{\hat{T}}{\varepsilon}\equiv\frac{1}{\varepsilon}\left\{1+\sum_{x=1}^{\infty}\frac{\lambda_X^x}{[x+\varepsilon][(x-1)+\varepsilon]\cdots[1+\varepsilon]}\right\}\sim O(\frac{1}{\varepsilon}) & x_0=0\\ \sum_{x=x_0}^{\infty}\frac{\lambda_X^{x-x_0}}{[x+\varepsilon][(x-1)+\varepsilon]\cdots[x_0+\varepsilon]}\sim O(1) & x_0>0\end{cases} &(43)\end{aligned}$$

When ε is small, the first term $\frac{\hat{T}}{\varepsilon}$ in Eq. (43) will be much larger than the lower terms. Similarly for the second moment we can derive from (39) that

$$\begin{aligned}\left\langle T_{x_0\to -1}^2\right\rangle &= \phi(x_0-1)\sum_{x=x_0}^{\infty}\frac{2\left\langle T_{x\to -1}\right\rangle}{\lambda_X\phi(x)} &(44)\\ &= \begin{cases} 2\frac{\hat{T}^2}{\varepsilon^2}+O\left(\frac{1}{\varepsilon}\right) & x_0=0\\ O(1) & x_0>0\end{cases} &(45)\end{aligned}$$

where the $x_0 = 0$ case follows since

$$\begin{aligned}\left\langle T_{0\to -1}^2\right\rangle &= \sum_{x=0}^{\infty}2\left[\frac{\hat{T}}{\varepsilon}+O(1)\right]\cdot\frac{1}{\lambda_X\phi(x)}\\ &= 2\frac{\hat{T}^2}{\varepsilon^2}+\sum_{x=0}^{\infty}\frac{1}{\lambda_X\phi(x)}\cdot O(1)\\ &= 2\frac{\hat{T}^2}{\varepsilon^2}+O\left(\frac{1}{\varepsilon}\right) &(46)\end{aligned}$$

Proceeding iteratively, we have for (40)

$$\begin{aligned}\left\langle T_{x_0\to -1}^n\right\rangle &= \phi(x_0-1)\sum_{x=x_0}^{\infty}\frac{n\left\langle T_{x\to -1}^{n-1}\right\rangle}{\lambda_X\phi(x)} &(47)\\ &= \begin{cases} n!\frac{\hat{T}^n}{\varepsilon^n}+O\left(\frac{1}{\varepsilon^{n-1}}\right) & x_0=0\\ O(1) & x_0>0\end{cases} &(48)\end{aligned}$$

where the $x_0 = 0$ case arises because

$$\begin{aligned}\left\langle T_{0\to -1}^n\right\rangle &= \sum_{x=0}^{\infty}n\left[(n-1)!\frac{\hat{T}^{n-1}}{\varepsilon^{n-1}}+O\left(\frac{1}{\varepsilon^{n-2}}\right)\right]\cdot\frac{1}{\lambda_X\phi(x)}\\ &= n!\frac{\hat{T}^n}{\varepsilon^n}+\sum_{x=0}^{\infty}\frac{1}{\lambda_X\phi(x)}\cdot O\left(\frac{1}{\varepsilon^{n-2}}\right)\\ &= n!\frac{\hat{T}^n}{\varepsilon^n}+O\left(\frac{1}{\varepsilon^{n-1}}\right) &(49)\end{aligned}$$

As $\varepsilon\to 0$, the first term accounts for more and more of the moment $\left\langle T_{0\to -1}^n\right\rangle$. We thus conclude the first passage time is asymptotically exponentially distributed with mean $\hat{T}/\varepsilon$.

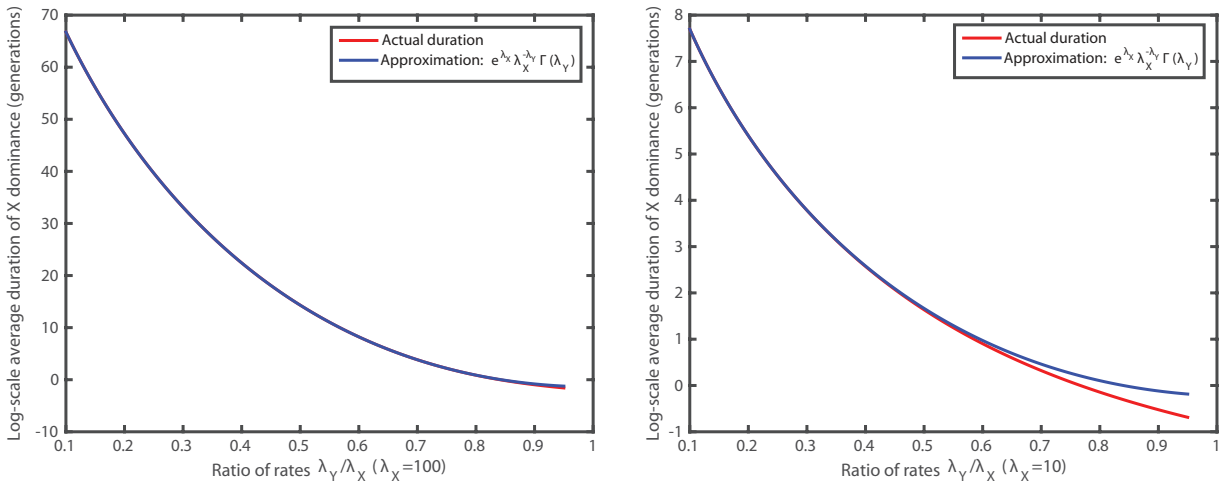


Figure 7: **Approximation for SinR dominance period durations.** In the fast complexing limit, Eq. 50 yields an estimate of the average duration of dominance of the majority species X. The figure compares the bound to the actual dominance periods in two regimes as the ratio of production rates λ_Y/λ_X is changed.

5.2 Periods of majority dominance are long and highly sensitive to the ratio of production rates λ_Y/λ_X

As in the case of the minority species, the estimates above for the first passage time can also be used to investigate the properties of the majority dominant state.

$$\begin{aligned}\langle T_{0\to -1}\rangle &= \frac{1}{\lambda_X}\sum_{x=0}^{\infty}\frac{\lambda_X^x}{[x+\lambda_Y][x-1+\lambda_Y]\cdots[1+\lambda_Y]}\frac{\lambda_X}{\lambda_Y}\\ &= e^{\lambda_X}(\lambda_X)^{-\lambda_Y}\,\gamma(\lambda_Y,\lambda_X)\\ &= e^{\lambda_X}(\lambda_X)^{-\lambda_Y}\,\Gamma(\lambda_Y) - e^{\lambda_X}(\lambda_X)^{-\lambda_Y}(\lambda_X)^{\lambda_Y-1}e^{-\lambda_X}\\ &\approx e^{\lambda_X}(\lambda_X)^{-\lambda_Y}\,\Gamma(\lambda_Y)\end{aligned} \tag{50}$$

Although better approximations are possible, it turns out that Eq. 50 is a good approximation for the duration of SinR dominance periods, as seen in Fig. 7. Note that dominance periods for the majority species are much longer than for the minority species, and that the duration is highly sensitive to the ratio of production rates λ_Y/λ_X.

5.3 The connection to experimentally measured distributions

The previous argument calculated the distribution of dominance period durations for the majority species SinR (X) and showed that they were exponentially distributed when switching is rare. In experiments, we instead measure the inter-pulse durations, i.e. the time from the beginning of one minority dominance period to the next. We next briefly illustrate that the distribution of these times should be essentially the same.

The key observation is that the minority dominance periods are short and narrowly distributed, as proven above. The intervals measured in the interpulse distributions are thus the sum of an exponentially distributed period in which SinR dominates, and a short, narrowly distributed period in which SinI dominates. We can thus approximate it is a constant. The intervals we measure experimentally (which we see are exponentially distributed) then take the form $t = t' + t_0$, where t is the observed interpulse duration, t is the "true" SinR dominance period, and t_0 is a constant representing the period of SinI dominance. Because the exponential distribution is memoryless, the "distribution minus a constant" has the same distribution—and

thus the same mean—as the original random variable. Thus if we observe an exponential $P(t) = \lambda e^{-\lambda t}$, the random variable $t' = t - t_0$, $t' \geq 0$ has a distribution

$$P(t') = P_t(t_0 + t' | t' \geq 0) = \frac{\lambda e^{-\lambda(t'+t_0)}}{\int_{t=t_0}^{\infty} \lambda e^{-\lambda t} dt} = \frac{\lambda e^{-\lambda(t'+t_0)}}{e^{-\lambda t_0}} = \lambda e^{-\lambda t'} \tag{51}$$

Hence the SinR dominance periods (t') follow the same exponential distribution as the observed interpulse durations (t).

6 SlrR can prolong dominance periods

The previous sections highlight the importance of the parameter λ_Y/λ_X in understanding the dynamics of SinI and SinR. Comparing Fig. 3 to Fig. 7, we see that the average duration of SinI dominance is extraordinarily insensitive (in absolute terms) to this parameter, while the period of SinR dominance is conversely extraordinarily sensitive. The net effect is that setting the relative production rates sets the frequency of pulsing, and once this is set the resulting pulses are all approximately impulsive in nature. Bursting can expand the duration of a period of SinI dominance, but this is opposed by the exponential nature of dilution/degradation, meaning that there are practical limits to how long a pulse can be made using the SinI and SinR circuit alone. Bursting can, however, control how narrowly distributed dominance periods are, as seen above. Below we demonstrate that the addition of SlrR to the circuit provides a vital "third knob" that allows the periods of SinI dominance to be expanded.

6.1 A model with SlrR and a reporter

To model the effect of SlrR, we introduce a third species Z to the model of Eq. 27:

$$\begin{aligned} X \xrightarrow{P_{N_a}(a)\cdot\lambda_X} X + a,\ \forall a \geq 0 \qquad Y \xrightarrow{P_{N_b}(b)\cdot\lambda_Y} Y + b,\ \forall b \geq 0 \\ X \xrightarrow{X/\tau_X} X - 1 \qquad Y \xrightarrow{Y/\tau_Y} Y - 1 \\ (X, Y) \xrightarrow{cXY} (X - 1, Y - 1) \\ Z \xrightarrow{f(X)} Z + 1 \qquad Z \xrightarrow{Z/\tau_Z} Z - 1 \\ (X, Z) \xrightarrow{dXZ} (X - 1, Z - 1) \end{aligned} \tag{52}$$

In simulations we assume that $\tau_X = \tau_Y = \tau_Z = \tau := 1$ (proteins are diluted by cell division), c and d are both very large (fast complexing), both binding reactions are irreversible, and that SinR regulates SlrR via a Hill function dependence: $f(X) = \frac{\lambda_Z K_d}{K_d + X^h}$. We additionally include a reporter species R, representing a repressed target of SinR:

$$R \xrightarrow{P_{N_r}(r)\cdot g(X)} R + r, \qquad R \xrightarrow{R/\tau} R - 1$$

where repression follows Hill-type dynamics $g(X) = \frac{\lambda_R}{1+(X/K_r)^n}$. In this section we will define dominance periods through the expression of this reporter, as it provides a simple synthesis of SinI and SlrR's effects.

6.2 Prolonged ON durations in the presence of SlrR

With two antagonists of SinR present, it is possible to use either to control the duration of ON durations (which now correspond to periods when the pool of free SinI + SlrR exceeds that of SinR or, roughly equivalently, when the reporter gene turns ON). Figure 8 compares these two strategies. In both figures, SinR is produced at a set level ($\lambda_X = 40$, and geometric bursts of average size $N_a = 10$). In the left panel, we sweep SinI by changing its production rate (λ_Y between 20 and 40, geometric bursts of average size $N_b = 10$). In the right panel, we keep SinI's production constant ($\lambda_Y = 20$, $N_b = 10$) and instead increase ON durations

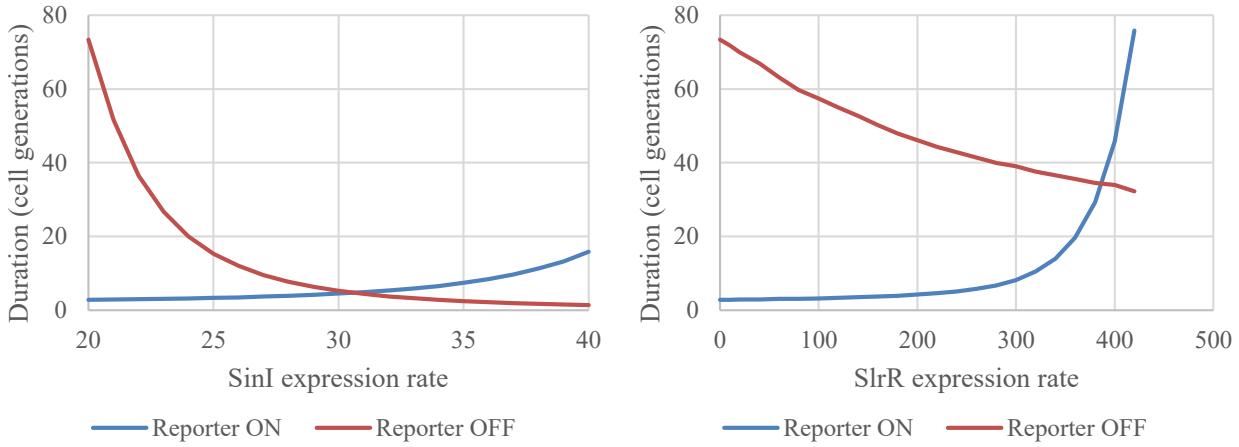


Figure 8: Comparison of effect of increasing SinI production (left) and increasing SlrR for fixed SinI production (right) on reporter ON and OFF durations. Production rates are described in the text. Other parameters are: $c = d = 10000$, $K_r = 5$, $K_d = 10^4$,$h = n = 4$,$\lambda_R = 5$, $N_r = 5$, and all lifetimes ($\tau's$) are set to 1, representing common removal by dilution/degradation.

by adding SlrR to the system (sweeping λ_Z from 0 to 440). As seen in the figures, this allows for much greater variation in ON durations with relatively little effect on OFF durations, as opposed to the tight correlation between the two in the circuit without SinI. By altering properties of the feedback loop, these differences in behavior can be further tuned.

Importantly, the addition of SlrR does not alter the main properties of the circuit, i.e. the memoryless production of stereotyped pulses. Figure 9 compares the distributions of reportor ON and OFF durations in circuits with and without SlrR.

7 Detailed modeling of the SinI-SinR interaction on downstream targets yields no qualitative differences in behavior

In this section, we briefly discuss a more mechanistically complex model of the interaction between SinI and SinR. The reactions below describe a system in which SinR (X) and SinI (Y) are produced in geometrically-distributed bursts and removed by dilution/degradation, as in the main SI. The next series of reactions explicitly model SinR dimerization (species DX) and tetramerization (species TX). We model the effects of SinI on SinR as targeting the active tetramer population — that is, SinR dimers either merge to form tetramers, which are capable of binding DNA, or are dismantled by permanent association with SinI to form the species XY, which is then removed by dilution/degradation.

The target of SinR (R) is modeled as having two binding sites. The number bound is kept track of by the variable B, which has a maximum value of 2. Transcription of the reporter takes place only when both sites are unbound, and is made in geometrically-distributed bursts. The total model thus comprises the following reactions:

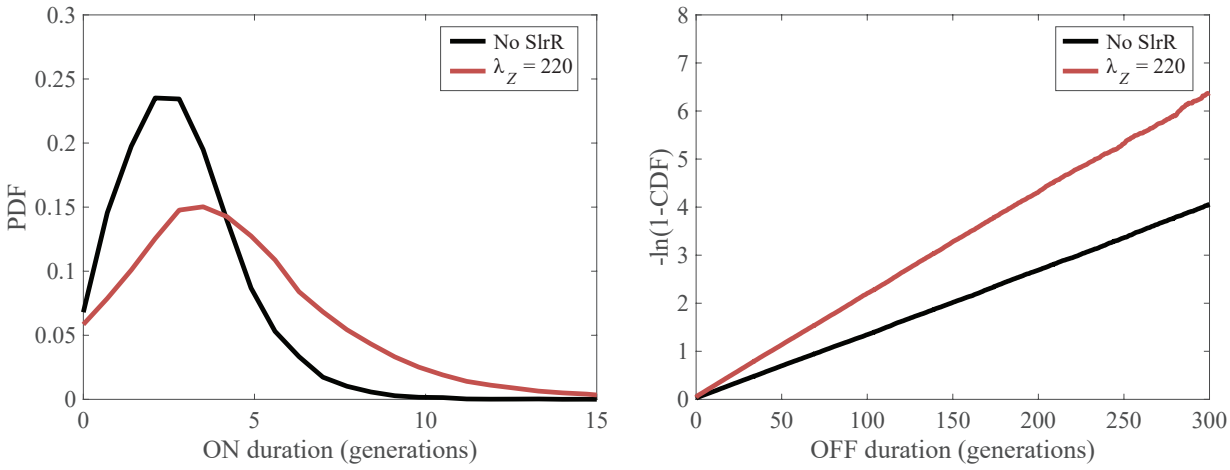


Figure 9: Example distributions of reporter ON (left panel) and OFF (right panel) durations in the circuit with SlrR (red lines) and without (black lines). Parameters are as in Figure 8 with $\lambda_Z = 220$.

$$X \xrightarrow{P_{N_a}(a)\cdot\lambda_X} X + a,\ \forall a \geq 0 \qquad Y \xrightarrow{P_{N_b}(b)\cdot\lambda_Y} Y + b,\ \forall b \geq 0$$

$$X \xrightarrow{X/\tau_X} X - 1 \qquad Y \xrightarrow{Y/\tau_Y} Y - 1$$

$$(X, DX) \xrightarrow{\lambda_{DX}\cdot X\cdot(X-1)} (X - 2, DX + 1)$$

$$(DX, TX) \xrightarrow{\lambda_{TX}\cdot DX\cdot(DX-1)} (DX - 2, TX + 1)$$

$$(DX, TX) \xrightarrow{d_{TX}\cdot TX} (DX + 2, TX - 1)$$

$$DX \xrightarrow{DX/\tau_{DX}} DX - 1 \qquad TX \xrightarrow{TX/\tau_{TX}} TX - 1$$

$$(DX, Y, XY) \xrightarrow{c\cdot DX\cdot Y(Y-1)} (DX - 1, Y - 2, XY + 2) \qquad XY \xrightarrow{XY/\tau_{XY}} XY - 1$$

$$(B, TX) \xrightarrow{\lambda_B\cdot(2-B)\cdot TX} (B + 1, TX - 1)$$

$$(B, TX) \xrightarrow{d_B\cdot B} (B - 1, TX + 1)$$

$$R \xrightarrow{(1-B)\cdot(2-B)\cdot\lambda_R\cdot P_{N_r}(r)} R + r,\ \forall r \geq 0 \qquad R \xrightarrow{R/\tau_R} R - 1$$

where λ_X, λ_Y are the production rates of SinR and SinI mRNA molecules; τ_X, τ_Y, τ_{DX}, τ_{TX}, τ_{XY} are the half-lives of X, Y, DX, TX, and XY proteins, c is the rate constant for complex formation, and λ_{DX}, λ_{TX} and d_{TX} are the rate constants controlling SinR dimer/tetramer formation. We assume all three proteins are subject to the same degradation pathway and/or dilution through cell growth, and thus have the same half-lives. For convenience we thus set $\tau_X = \tau_Y = \tau_{XY} = \tau_{DX} = \tau_{TX} = \tau_R = 1$. Production of X,Y and R proteins are produced in bursts: a, b, and r are geometrically distributed random variables with average burst sizes N_a, N_b, and N_r.

Many of the rate constants are not easily measurable, so we use a combination of constraints from experimental data and reasonable assumptions:

- Though we do not know the exact abundance of each form of SinI and SinR, we can use rough literature estimates for the total abundance of all forms. We first constrain $\langle \textbf{SinR total} \rangle = \langle X + 2 \cdot DX + 4 \cdot TX + XY \rangle \approx 400$, in line with mass spec estimates from growth in rich media [30]. SinI is low abundance and difficult to measure directly. We approximate its average copy number as $\langle \textbf{SinI total} \rangle = \langle Y + XY \rangle \approx 100$

based on our HALO-tagging experiments, though this estimate is rough since the two strains are constructed differently.

- SinI affinity for SinR is measured to be in the nanomolar range [31], so we model this reaction as irreversible.
- SinR is essentially completely tetrameric in solution (Anthony Wilkinson, personal communication, and by similarity to the cI-Cro system [32]), so adjust λ_{DX}, λ_{TX}, and d_{TX} to favor that ($\lambda_{DX} = 10000$, $\lambda_{TX} = 1$, $d_{TX} = 100$). We note this is in contradiction to a report in the literature of a K_d of 6.7 μM for SinR tetramerization ([33]). This number was measured under non-physiological conditions and, were it true, it would be expected that almost no tetramers would ever be present in the cell.
- Likewise, from our HALO-tagging experiments, we know approximate CVs for the these totals, so we similarly constrain $CV_{\text{SinR total}} \approx 0.27$ and $CV_{\text{SinI total}} \approx 0.34$.
- From footprinting and gel-shift experiments, we observe almost complete protection of endogenous SinR targets with 40 nM SinR ([34], [35]). We thus set the affinity of SinR tetramers for binding sites to be 40 nM to be conservative ($\lambda_B = 1$, $d_B = K_{d,B}\lambda_B = 40$). We note that there are two papers in the literature claiming affinities of SinR for DNA in the 350 nM range ([36], [31]). Both papers measure SinR binding to a short synthetic oligo sequence and not to natural promoters. Moreover, the measured numbers are inconsistent with all copy number estimates of SinR in individual cells, as at this affinity very little SinR would ever be bound to DNA.
- We assume all proteins are subject to removal by dilution/degradation with the same rate constants (all τ's set to 1).

The constraints still leave a lot of parametric flexibility, including the complete dynamics of SinI/SinR production. One example sets the average bursting rate of SinI to 10 ($\lambda_Y = 10$) with average burst size 10 ($N_b = \langle b \rangle = 10$) and average burst rate of SinR to 16 ($\lambda_R = 16$) with average burst size 25 ($N_a = \langle a \rangle = 25$). These dynamics lead to average abundances and CVs that match the estimates above. The second area of flexibility is in reporter expression. We thus provide a handful of examples below.

In most of the SI we ignored this level of complexity and used simpler models describing only the interaction of SinI and SinR, as it is this interaction that gives rise to most of the behavior. To show that these simpler models accurately recapitulate the "more complete" model above, we simulated a reduced version in which explicit modeling of SinR oligomerization and DNA binding was ignored and replaced by Hill-type dynamics:

$$\text{reporter production rate} = \frac{\lambda_R}{1 + (X/K_d)^h}$$

By matching the production rates and equilbrium concentrations of SinR between these two models, we can produce a "matched" reduced model. The precise matching conditions are described in the next section below, but the results boil down to doubling λ_R and setting $K_d = 249$. Simulations of both models are showing in Figure 10, showing excellent agreement in the qualitative behavior despite the simplifications.

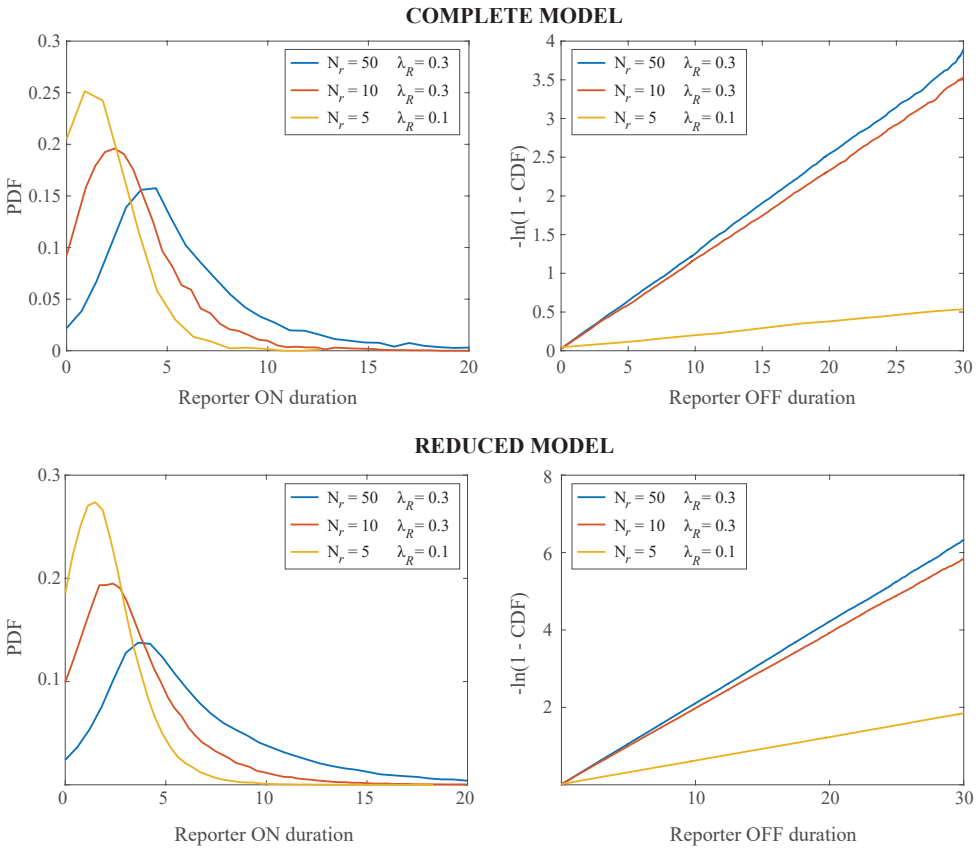


Figure 10: The figure shows the distributions of SinI dominance (left panels) and SinR dominance (right panels) as measured by expression of the SinR-controlled reporter R. The top panels are for the complete model described above, and the bottom panels are for the reduced model where SinR tetramerization and promoter binding have been approximated by Hill-function dynamics. The parameters of the reporter are indicated by the line color.

7.1 Deriving a matched, reduced model from the complete model

We will approximate that all of the X molecules are either in dimers or tetramers, as the equilibria strongly favor oligomerization of X. The key reactions in our "complete model" involving multimers of X are then:

$$\begin{aligned}
(DX, TX) &\xrightarrow{\lambda_{TX}\cdot DX\cdot(DX-1)} (DX-2, TX+1)\\
(DX, TX) &\xrightarrow{d_{TX}\cdot TX} (DX+2, TX-1)\\
(B, TX) &\xrightarrow{\lambda_B\cdot(2-B)\cdot TX} (B+1, TX-1)\\
(B, TX) &\xrightarrow{d_B\cdot B} (B-1, TX+1)\\
R \xrightarrow{(1-B)\cdot(2-B)\cdot\lambda_R\cdot P_{N_r}(r)} R+r,\ \forall r\geq 0 \qquad & R \xrightarrow{R/\tau_R} R-1
\end{aligned}$$

We approximate the two tetramer binding sites as one with double the production rate, leading to the modified reactions:

$$\begin{aligned}
(B, TX) &\xrightarrow{2\lambda_B\cdot(1-B)\cdot TX} (B+1, TX-1)\\
(B, TX) &\xrightarrow{d_B\cdot B} (B-1, TX+1)\\
R \xrightarrow{2(1-B)\cdot\lambda_R\cdot P_{N_r}(r)\equiv v_R\cdot P_{N_r}} R+r,\ \forall r\geq 0 \qquad & R \xrightarrow{R/\tau_R} R-1
\end{aligned}$$

The mean dynamics of dimers, tetramers, and DNA binding then follow:

$$\begin{aligned}
\frac{dDX}{dt} &= -2\lambda_{TX}\cdot DX^2 + 2d_{TX}\cdot TX\\
\frac{dTX}{dt} &= \lambda_{TX}\cdot DX^2 - d_{TX}\cdot TX - 2\lambda_B\cdot(1-B)\cdot TX + d_B\cdot B\\
\frac{dB}{dt} &= 2\lambda_B\cdot(1-B)\cdot TX - d_B\cdot B
\end{aligned}$$

In equilibrium, $\frac{dDX}{dt} = \frac{dTX}{dt} = \frac{dB}{dt} = 0$, so

$$\begin{aligned}
-\lambda_{TX}\cdot DX^2 + d_{TX}\cdot TX &= 0\\
2\lambda_B\cdot(1-B)\cdot TX - d_B\cdot B &= 0
\end{aligned}$$

We can then calculate equilibrium values for the concentrations of the various species, the bound fraction, and the production rate of the of the reporter:

$$\begin{aligned}
TX &= DX^2/K_{d,TX}\\
B &= \frac{2\lambda_B TX}{d_B + 2\lambda_B TX}\\
v_R &= 2\lambda_R\cdot(1-B) = \frac{2\lambda_R}{1+TX/(K_{d,B}/2)} = \frac{2\lambda_R}{1+DX^2/(K_{d,TX}\cdot K_{d,B}/2)}
\end{aligned} \tag{53}$$

In an equivalent "reduced" model, we ignore explict modeling of the dimer and tetramer species and model binding simply as 4 X molecules coming together on a single site on DNA:

$$\begin{aligned}
(X, B) &\xrightarrow{\lambda_B\cdot X(X-1)(X-2)(X-3)(1-B)} (X-4, B+1)\\
(X, B) &\xrightarrow{d_B\cdot B} (X+4, B-1)\\
R \xrightarrow{(1-B)\cdot\lambda'_R\cdot P_{N_r}(r)\equiv v_R\cdot P_{N_r}} R+r,\ \forall r\geq 0 \qquad & R \xrightarrow{R/\tau_R} R-1
\end{aligned}$$

The equilibrium reactions in this case follow:

$$\frac{dX}{dt} = -4\lambda_B \cdot X^4 \cdot (1-B) + 4d_B \cdot B$$
$$\frac{dB}{dt} = \lambda_B \cdot X^4 \cdot (1-B) - d_B \cdot B$$

And in equilibrium, $\frac{dX}{dt} = \frac{dB}{dt} = 0$, we have

$$\begin{aligned} \lambda_B \cdot X^4 \cdot (1-B) &= d_B \cdot B \\ B &= \frac{\lambda_B \cdot X^4}{d_B + \lambda_B \cdot X^4} \\ v_R = \lambda'_R \cdot (1-B) &= \frac{\lambda'_R}{1 + X^4/K'^4_{d,B}} = \frac{\lambda'_R}{1 + (X/K_d)^4} \end{aligned} \tag{54}$$

By comparing Eqs. 53 with Eqs. 54, we can match production rates

$$v_R = \frac{\lambda'_R}{1 + (X/K_d)^4} = \frac{2\lambda_R}{1 + DX^2/(K_{d,TX} \cdot K_{d,B}/2)} \tag{55}$$

Since we assume in the complete model that complexing is fast (so that all X is in either dimers or tetramers), we then can dictate that the total equilibrium average amounts of X between the complete (LHS) and reduced (RHS) models be the same:

$$\begin{aligned} 2 \cdot DX + 4 \cdot TX &= X \\ 2 \cdot DX + 4 \cdot DX^2/K_{d,TX} &= X \qquad \text{(equilibrium between } DX \text{ and } TX\text{)} \\ &\approx \lambda_X \cdot N_a - \lambda_Y \cdot N_b \end{aligned}$$

The final approximation follows from the dynamics of X production and removal by Y (assuming complexing by SinI is dominates over dilution). By solving the equation using the parameters in the previous section, we have $X \approx 300$, $DX \approx 65$. From Eq. (55), we obtain

$$\begin{aligned} \lambda'_R &= 2\lambda_R \\ K_d &= (K'_{d,B})^{1/4} = (X^4/DX^2 \cdot K_{d,TX} \cdot K_{d,B}/2)^{1/4} \approx 249 \end{aligned}$$

8 Expected autocorrelation in experimental reporters

In the main text and above, we argue that stochastic antagonism between SinI and SinR gives rise to dynamics in which the minority species comes to dominate for short periods of time at memoryless intervals. However, because these dominance periods are so short, they cannot be observed directly, though we can look for properties such as memorylessness and timing that are expected to follow from the model. We next consider another prediction of the model that can be investigated experimentally: the expected autocorrelation for a reporter that is controlled by SinR (as plotted in Fig. S3C of the main text).

We must thus model the production and maturation of a fluorescent reporter under SinR control. We will denote immature (dark) fluorescent proteins by p and mature fluorescent protens by q. When durations of SinI dominance are short relative to the time between successive events, we can approximate the birth dynamics of p as occurring in large bursts (of size b) with constant probability in time (with rate λ, set by the relative balance of SinI and SinR). The protein then matures with first order kinetics (rate constant γ), and both mature and immature proteins are degraded/diluted through growth (occurring exponentially with rate β). This leads to the scheme

$$p \xrightarrow{\lambda} p + b \qquad (p, q) \xrightarrow{\gamma p} (p-1, q+1)$$
$$p \xrightarrow{\beta p} p - 1 \qquad q \xrightarrow{\beta q} q - 1$$

We assume that p and q are sufficiently high abundance that low copy fluctuations are negligible, and the only source of noise in the system is production bursts (*i.e.* $\frac{1}{\langle p \rangle} \approx 0$, $\frac{1}{\langle q \rangle} \approx 0$, but $\frac{b}{\langle p \rangle}$ is non-negligible). In this case the steady state normalized covariances $\eta_{pq} = \frac{\mathrm{Cov}(p,q)}{\langle p \rangle \langle q \rangle}$,$\eta_{qq} = \sigma_q^2 / \langle q \rangle^2$, and $\eta_{pp} = \sigma_p^2 / \langle p \rangle^2$ using the fluctuation-dissipation relation (FDR) [37]

$$M\eta + \eta M^t = D$$

with $M = \begin{pmatrix} \beta + \gamma & 0 \\ -\beta & \beta \end{pmatrix}$, $\eta = \begin{pmatrix} \eta_{pp} & \eta_{pq} \\ \eta_{pq} & \eta_{qq} \end{pmatrix}$ (the normalized covariance matrix), and $D = \begin{pmatrix} \frac{2b(\beta+\gamma)}{\langle p \rangle} & 0 \\ 0 & 0 \end{pmatrix}$. The zeros in the diffusion matrix D are induced by the approximations mentioned above, and indicate that we consider only noise originating in p as it propagates through the system. The solution then yields

$$\eta_{pp} = \frac{b}{\langle p \rangle}$$
$$\eta_{pq} = \eta_{qq} = \frac{\beta}{(\gamma + 2\beta)} \frac{b}{\langle p \rangle}$$

A slight modification of the calculation in [38] then shows that the normalized matrix of autocorrelations ($R(\tau)$) follows

$$\frac{dR(\tau)}{dt} = -MR(\tau)$$

with solution given by the matrix exponential

$$R(\tau) = e^{-M\tau}\eta \quad .$$

After normalizing by η_{qq}^2, we thus obtain

$$\frac{R_{qq}(\tau)}{\eta_{qq}^2} = \frac{\beta}{\gamma} e^{-\beta\tau} - \frac{e^{-\beta\tau}}{\gamma}(\beta e^{-\gamma\tau} - \gamma) \quad .$$

Remarks

- Once the autocorrelation has been normalized to its initial value, the curve is parametrized by only the component stability (β), and the maturation rate of the fluorescent protein (γ). Empirically-determined values for these constants are used to plot the curves in the main text.
- The initiation rate of bursts (i.e. λ) does not enter into this equation: provided that bursts fire memorylessly in time, the temporal correlations will always follow the same form. The firing rate does of course change the moments, but these changes are scaled away by normalizing the autocorrelation to its initial value (i.e. η_{qq}^2).
- Consistency with this model indicates that production of the immature reporter can be treated as an essentially instantaneous process. This is consistent with the expectations derived from the simple stochastic antagonism model in which periods of SinI dominance are expected to be rare and very short. Correlations that decay more slowly than this expectation are suggestive of a long-lived variability upstream of the reporter itself. We in fact observe precisely this effect when we add SlrR to our reconstitution system. In that circuit, the periods of SinI/SlrR dominance have substantial duration, creating longer-lived temporal correlations in the reporter.

9 Inferring properties of stochastic antagonism from reporter traces

In the main text, we argued that the observed pulses of gene expression in *B. subtilis* and the reconstituted *E. coli* circuit resulted from occasional dominance of SinI over SinR. As suggested by the mathematical

reasoning in previous sections of this document, these periods of dominance are expected to be very short—well below the time resolution of our experiments. While the statistics describing the periods of dominance of SinI over SinR are thus not directly accessible, we can make some inferences using the observed peaks and additional measurements of variability in GFP synthesis under saturating IPTG conditions. Specifically, we would like to determine whether the data are consistent with a model in which the periods of active reporter synthesis (*i.e.* SinI dominance) exhibit timing.

9.1 A simple model for reporter burst size

To make these inferences, we developed a simple model describing how bursts of reporter synthesis follow from the underlying dynamics of SinI and SinR. We assume that bursts arise from discrete periods of promoter activity, each of which has a random duration (T, with unknown statistics). During this period, reporter production is assumed to follow a non-homogeneous Poisson process with possibly changing rate $\lambda(t)$. The size of burst produced in a period of SinI dominance can then be written as

$$\begin{aligned} b &= \int_0^T \lambda(t')dt' \\ &= T\left(\frac{1}{T}\int_0^T \lambda(t')dt'\right) \\ &= T\bar{\lambda} \end{aligned}$$

where $\bar{\lambda}$ is the average production rate taken over the duration of SinI dominance (itself a random variable). We next make the simplifying (but reasonable, see below) assumption that these two random variables can be treated as independent of one another, so that the normalized variance in the burst size $\frac{\sigma_b^2}{\langle b\rangle^2}$ then follows:

$$\frac{\sigma_b^2}{\langle b\rangle^2} = \frac{\sigma_T^2}{\langle T\rangle^2} + \frac{\sigma_{\bar{\lambda}}^2}{\langle\bar{\lambda}\rangle^2} + \frac{\sigma_T^2}{\langle T\rangle^2}\frac{\sigma_{\bar{\lambda}}^2}{\langle\bar{\lambda}\rangle^2} \tag{56}$$

The assumption of independence amounts to a statement about the relative timescales of synthesis rate fluctuation and duration of SinI dominance. Specifically, we assume that the rate is effectively constant within a given period of SinI dominance, but may fluctuates between subsequent intervals. Given that we establish the periods of SinI dominance to be substantially shorter than a generation, we consider this to be a reasonable assumption as most determinants of gene expression variability are stable (e.g. sigma factors, ribosomes, polymerases, etc.), and will likely fluctuate with a characteristic timescale of at least one generation.

9.2 Inferring timing in the SinI dominance periods

Eq. 56 can be rearranged to yield an estimate of the variability in SinI dominance periods T:

$$CV_T^2 = \frac{\sigma_T^2}{\langle T\rangle^2} = \frac{\frac{\sigma_b^2}{\langle b\rangle^2} - \frac{\sigma_{\bar{\lambda}}^2}{\langle\bar{\lambda}\rangle^2}}{1 + \frac{\sigma_{\bar{\lambda}}^2}{\langle\bar{\lambda}\rangle^2}} \tag{57}$$

We functionally define timing by the observation of sub-exponential statistics in T, i.e. $CV_T < 1$. We thus had to experimentally estimate $\sigma_b^2/\langle b\rangle^2$ and $\sigma_{\bar{\lambda}}^2/\langle\bar{\lambda}\rangle^2$.

We estimate the variability in burst size ($\sigma_b^2/\langle b\rangle^2$) from the distribution of peak heights in the main text (Figure 3). By simulating the burst-maturation-dilution process modeled described in he previous section, we found that the peak height was proportional to the initial burst size. The normalized variance in the distribution of peak heights will therefore be the same as the normalized variance in burst sizes. Using this procedure, we estimate $\sigma_b^2/\langle b\rangle^2$ in the reconstitution strain (NDL-423) to be 2.04, 1.35, and 0.83 for 90 uM, 100 uM and 110 uM IPTG, respectively.

To determine $\sigma_{\bar{\lambda}}^2/\langle\bar{\lambda}\rangle^2$, we directly measured the variability in GFP production rate in the reconstitution strain (NDL-423) under saturating concentrations of IPTG. We assume, therefore, that during periods of

SinI dominance, the reporter is completely derepressed, and is therefore well-modeled by the maximally-induced promoter. The rates were calculated by measuring the total amount of GFP produced within each 8-minute interframe interval (see SI methods), yielding an estimate of $\sigma^2_{\bar{\lambda}}/\langle\bar{\lambda}\rangle^2 = 0.33$.

Plugging into Eq. 57, we arrive at estimates for CV_T of 1.13, 0.87 and 0.61 (for 90 uM, 100 uM and 110 uM IPTG). Thus as pulsing becomes stronger and more frequent, it also becomes more narrowly distributed.

Remarks

We argue that this procedure results in conservative estimates for the degree of timing associated with SinI dominance for several reasons:

1. Our estimate of the variability in the production rate is extremely conservative (i.e. small). The measured rate reflects substantially more time-averaging of the production rate than is possible during the abbreviated periods of SinI dominance. As Eq. 57 is decreasing in $\sigma^2_{\bar{\lambda}}/\langle\bar{\lambda}\rangle^2$, the true values of CV_T will be lower than our estimates.
2. We neglect low-copy noise in the production of the immature reporter. This effectively over-attributes variability in the observed burst size to variability in duration of SinI dominance, as our estimates of $\sigma^2_b/\langle b\rangle^2$ will be increased by the added noise.
3. We derive our estimate of $\lambda(t)$ fluctuations from a conditioned dataset in which observations derived from late stages of the cell cycle were discarded. This again leads to an underestimate of the variability in production rate giving rise to the bursts: these events occur at all cell cycle stages, and therefore will reflect variation in chromosomal locus copy number and total protein expression capacity of the cell.

Accordingly, even under assumptions that systematically overestimate the variability in SinI dominance duration, we find evidence of timing in the vigorous pulsing regimes (100 uM and 110 uM IPTG) consistent with the model-derived prediction that stochastic antagonism can give rise to timed periods of SinI dominance without the need for feedback.

10 Diffusion Approximations

Because the natural process we study is discrete, and the states with close to zero free molecules of SinI or SinR are very relevant, we use discrete point processes in the derivations. However, to help orient readers who are more familiar with diffusion approximations, where the discrete random events are captured by adding noise terms to otherwise deterministic equations, we here attempt to rephrase some basic results in that language. For our models, we did not find this approximation helpful to obtain intuitive analytical results, so this section is only intended to serve a pedagogical role.

10.1 Diffusion model

For the chemical reaction system where SinR (X) and SinI (Y) are produced one by one and $\lambda_X > \lambda_Y$

$$\begin{aligned} X &\xrightarrow{\lambda_X} X+1 & Y &\xrightarrow{\lambda_Y} Y+1 \\ X &\xrightarrow{X} X-1 & Y &\xrightarrow{Y} Y-1 \\ (X,Y) &\xrightarrow{cXY} (X-1,Y-1) \end{aligned} \tag{58}$$

we can use e.g. Gillespie's framework [39] to obtain a corresponding chemical Langevin equation from the reaction system:

$$\begin{cases} \dfrac{dX}{dt} = \lambda_X - X - cXY + \sqrt{\lambda_X}\cdot\xi_1 - \sqrt{X}\cdot\xi_3 - \sqrt{cXY}\cdot\xi_5 \\ \dfrac{dY}{dt} = \lambda_Y - Y - cXY + \sqrt{\lambda_Y}\cdot\xi_2 - \sqrt{Y}\cdot\xi_4 - \sqrt{cXY}\cdot\xi_5 \end{cases} \tag{59}$$

where the ξ_i $(i = 1, \ldots, 5)$ are temporally uncorrelated, statistically independent Gaussian white noise terms corresponding to the five reactions in (58). One approach to gain intuition from Eq. (59) is to examine the properties of the fixed points from the deterministic counterpart dynamics

$$\begin{cases} \dfrac{dX}{dt} = \lambda_X - X - cXY \\ \dfrac{dY}{dt} = \lambda_Y - Y - cXY \end{cases} \tag{60}$$

which has a single nonnegative fixed point (X^*, Y^*)

$$\begin{cases} X^* = \dfrac{\lambda_X - \lambda_Y - \frac{1}{c} + \sqrt{\left(\lambda_X - \lambda_Y - \frac{1}{c}\right) + \frac{4\lambda_X}{c}}}{2} \\ Y^* = \dfrac{\lambda_Y - \lambda_X - \frac{1}{c} + \sqrt{\left(\lambda_Y - \lambda_X - \frac{1}{c}\right) + \frac{4\lambda_Y}{c}}}{2} \end{cases} \tag{61}$$

The eigenvalues of the Jacobian matrix at this fixed point are both negative $\left(-1, -\sqrt{4c\lambda_X + (c\lambda_X - c\lambda_Y - 1)^2}\right)$, hence the fixed point is stable. In the fast complexing region $c \to \infty$, $(X^*, Y^*) \to (\lambda_X - \lambda_Y, 0)$, the number of Y will go to zero quickly, and we expect the diffusion approximation to perform poorly since it can go to negative values, and correction methods that add reflecting boundaries can in turn introduce undesirable artifacts [40].

10.2 Diffusion approximation for dominance periods of the minority species

Analogously to the results in Section 3 and assuming high rates of complexing reaction, we can consider the first passage time from b to 0 SinI molecules to understand dominance periods of the minority species:

$$Y \xrightarrow{\lambda_X + Y} Y - 1 \tag{62}$$

This can be approximated by a Langevin equation

$$\frac{dY}{dt} = A(Y) - \sqrt{B(Y)} \cdot \xi(t) = -(\lambda_X + Y) - \sqrt{\lambda_X + Y} \cdot \xi(t) \tag{63}$$

Using the backward Fokker-Planck equation [28], we know the first passage time $\langle T_{b \to 0} \rangle$ starting from b with an absorbing boundary at 0 and a reflecting boundary at infinity is

$$\langle T_{b \to 0} \rangle = 2 \int_0^b \frac{dz}{\psi(z)} \int_z^\infty \frac{\psi(Y)}{B(Y)} dY \tag{64}$$

where

$$\psi(Y) = \exp\left[\int_0^Y \frac{2A(z)}{B(z)} dz\right] = e^{-2Y} \tag{65}$$

Therefore, we know

$$\begin{aligned} \langle T_{b \to 0} \rangle &= 2 \int_0^b \frac{dz}{e^{-2z}} \int_z^\infty \frac{e^{-2Y}}{\lambda_X + Y} dY \\ &= \log\left(1 + \frac{b}{\lambda_X}\right) + e^{2(\lambda_X + b)} \operatorname{Ei}[2(\lambda_X + b)] - e^{2\lambda_X} \operatorname{Ei}[2\lambda_X] \end{aligned} \tag{66}$$

where $\operatorname{Ei}[x] = \int_x^\infty e^{-t}/t \, dt$ is the exponential integral function, which is is similar to the exact result

$$\langle T'_{b \to 0} \rangle = \sum_{Y=1}^{b} \frac{1}{\lambda_X + Y} \approx \log\left(1 + \frac{b}{\lambda_X}\right) \tag{67}$$

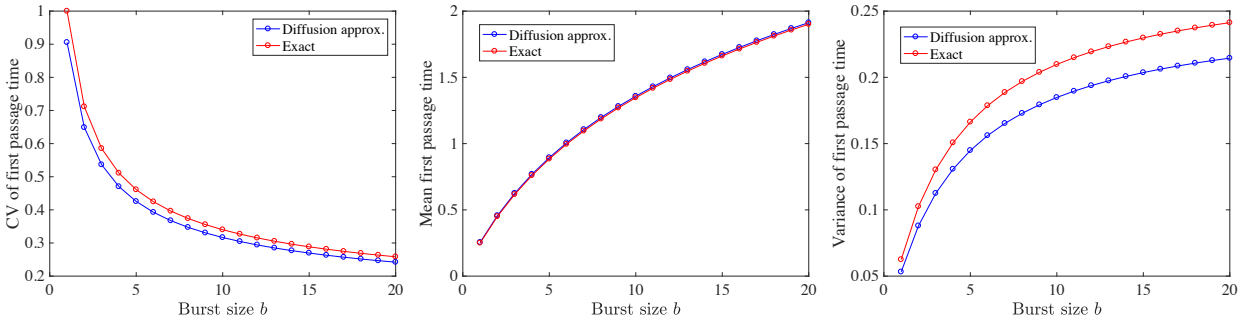


Figure 11: Comparison of the results under diffusion approximation with the exact results for fixed burst size b. Note that the $\lambda_X = 3$.

We can similarly obtain the second order moment of $T_{b\to 0}$ and its variance

$$\mathrm{Var}[T_{b\to 0}] = 4\int_0^b \frac{dz}{e^{-2z}} \int_z^\infty \frac{e^{-2Y}\left\langle T_{Y\to 0}\right\rangle}{\lambda_X + Y} dY - \left\langle T_{b\to 0}\right\rangle^2 \tag{68}$$

However, for the variance this integral seems if anything more difficult to calculate than the sum of the corresponding exact result

$$\mathrm{Var}[T'_{b\to 0}] = \sum_{Y=1}^{b} \frac{1}{(\lambda_X + Y)^2} \approx \frac{1}{\lambda_X} - \frac{1}{\lambda_X + b} \tag{69}$$

References

[28] C. W. Gardiner, *Handbook of Stochastic Methods: For Physics, Chemistry and the Natural Sciences* (Springer-Verlag, Berlin, 2004), third edn.

[29] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (Elsevier, Singapore, 2007), third edn.

[30] J. Muntel, *et al.*, *Molecular & Cellular Proteomics* **13**, 1008 (2014).

[31] J. A. Newman, C. Rodrigues, R. J. Lewis, *Journal of Biological Chemistry* **288**, 10766 (2013).

[32] R. J. Lewis, J. A. Brannigan, W. A. Offen, I. Smith, A. J. Wilkinson, *Journal of Molecular Biology* **283**, 907 (1998).

[33] D. J. Scott, *et al.*, *Journal of Molecular Biology* **293**, 997 (1999).

[34] F. Chu, D. B. Kearns, S. S. Branda, R. Kolter, R. Losick, *Molecular Microbiology* **59**, 1216 (2006).

[35] D. B. Kearns, F. Chu, S. S. Branda, R. Kolter, R. Losick, *Molecular Microbiology* **55**, 739 (2005).

[36] V. L. Colledge, *et al.*, *Journal of Molecular Biology* **411**, 597 (2011).

[37] J. Paulsson, *Nature* **427**, 415 (2004).

[38] I. Lestas, J. Paulsson, N. E. Ross, G. Vinnicombe, *IEEE Transactions on Automatic Control* **53**, 189 (2008).

[39] D. T. Gillespie, *The Journal of Chemical Physics* **113**, 297 (2000).

[40] D. Schnoerr, G. Sanguinetti, R. Grima, *The Journal of Chemical Physics* **141**, 024103 (2014).