Supplemental material for the paper: Noise analysis of genome-scale protein synthesis using a discrete computational model of translation

Julien Racle[a,b,c], Adam Jan Stefaniuk[a,d] and Vassily Hatzimanikatis[*,a,b]

[a]Laboratory of Computational Systems Biotechnology, École Polytechnique Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[b]Swiss Institute of Bioinformatics (SIB), CH-1015 Lausanne, Switzerland
[c]Current address: Computational Cancer Biology Lab, Ludwig Center for Cancer Research, University of Lausanne, CH-1066 Epalinges, Switzerland
[d]Current address: Via Loreto 24, CH-6900 Lugano, Switzerland

[*]Corresponding author: vassily.hatzimanikatis@epfl.ch

Supplemental Figures

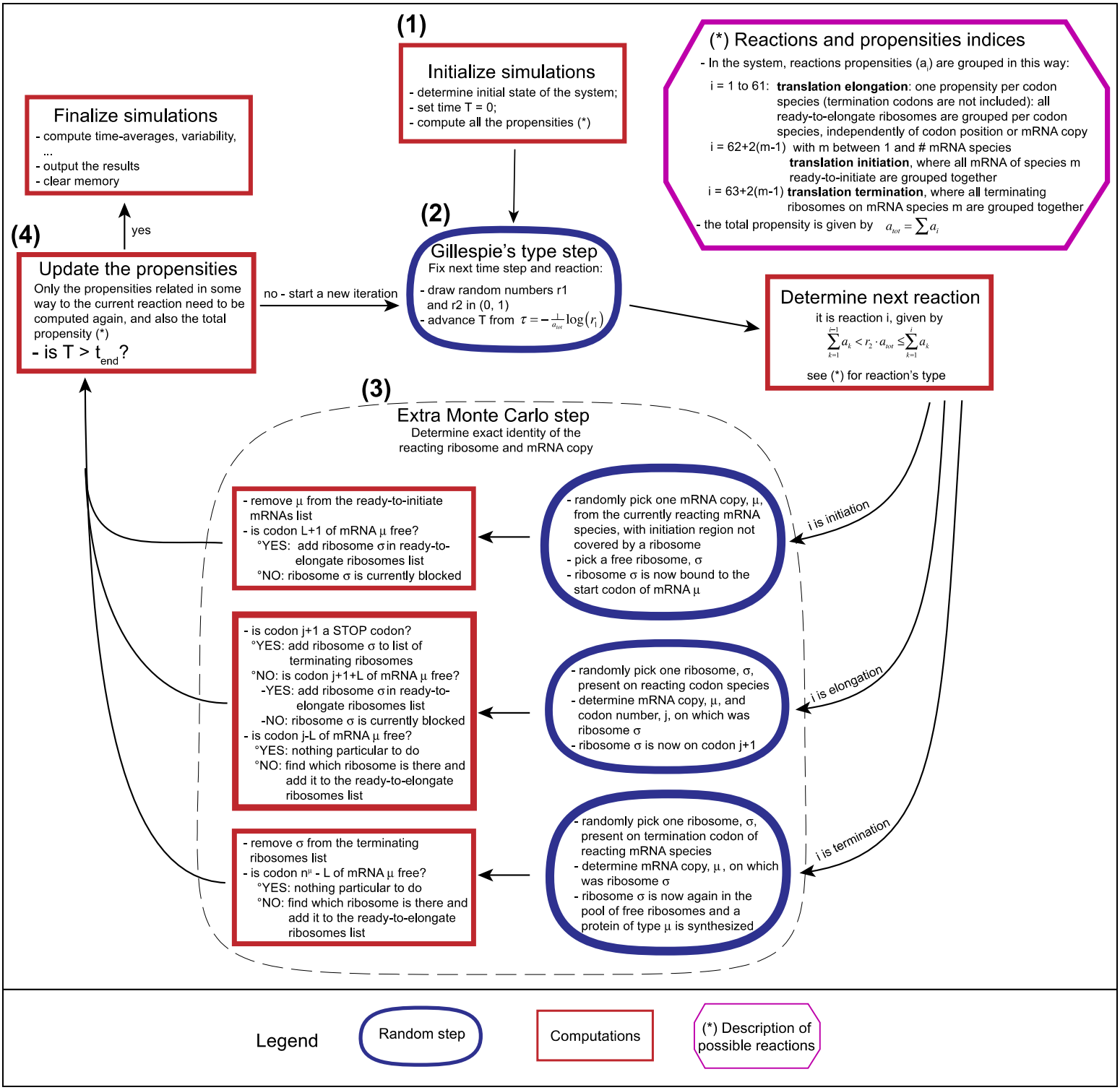


FIGURE S1 Flowchart of the stochastic algorithm used in the simulations.
The numbers 1-4 in parenthesis refer to descriptions in the text.

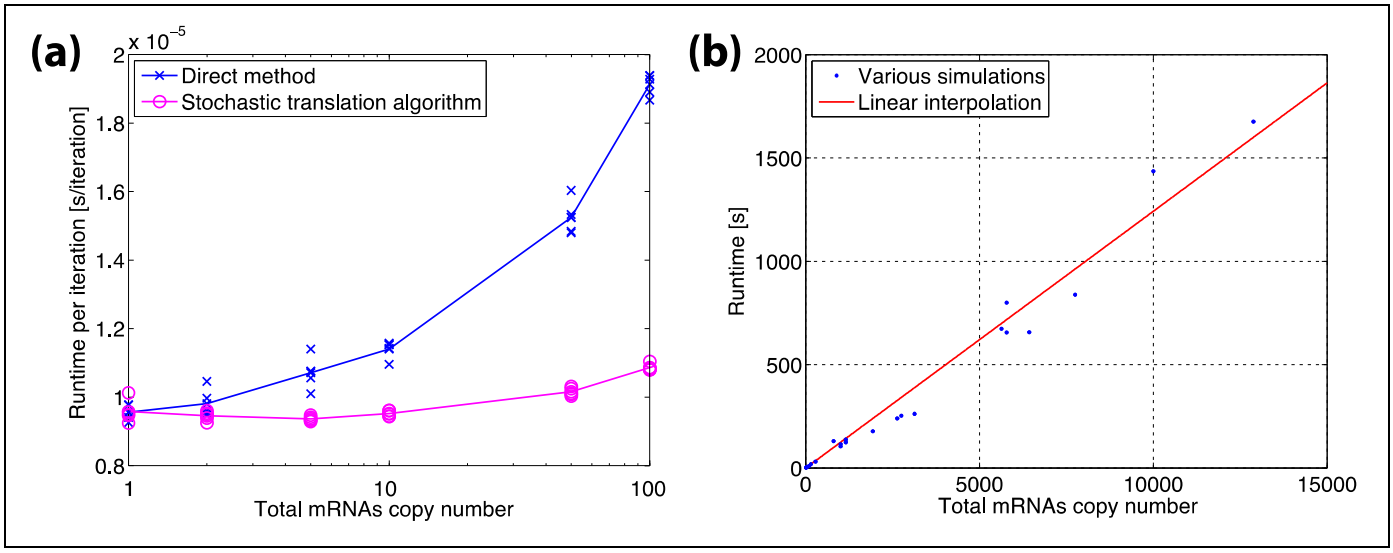


FIGURE S2 Runtime of the simulations. (a) Comparison of the runtime using Gillespie's optimized direct method or our stochastic translation algorithm - these simulations were implemented in MATLAB. Five simulations with randomly selected genes were performed for each value of total mRNA copy number. **(b)** The runtime with our stochastic translation algorithm, for simulations optimized in C++, scales only linearly with the total number of mRNA copies present. (The coefficient of determination R^2 of the linear fitting is 0.97). The simulations were performed on Mac Pro computer, with a 2 x 2.93 GHz Quad-Core Intel Xeon processor, on a C++ implementation of the algorithm that was not parallelized. The final time in these simulations was taken as 1000 seconds.

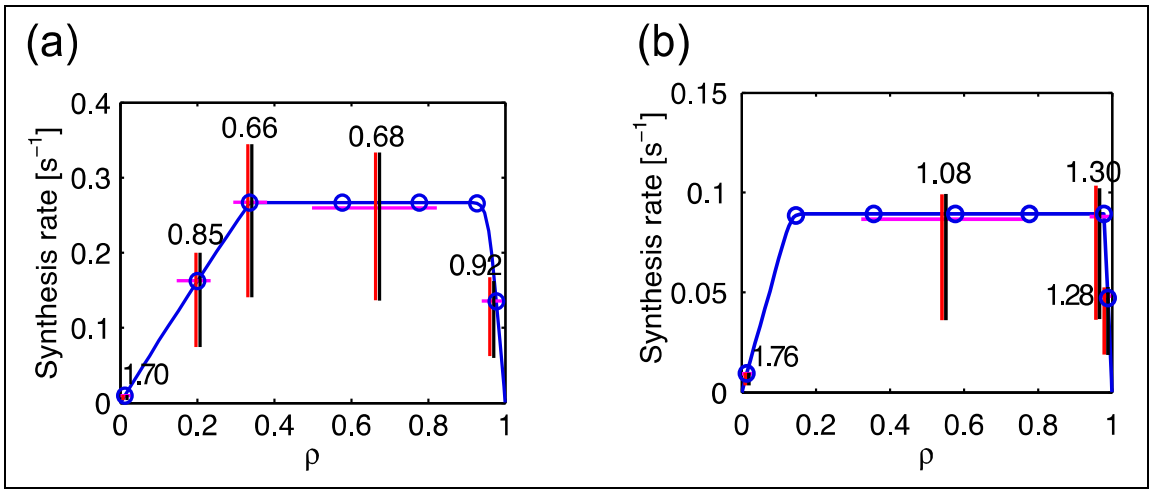


FIGURE S3 Synthesis rate profiles for suboptimal conditions. Similar profiles to Fig. 3 are shown but for conditions of initiation/termination rate constants that are suboptimal. See Fig. 3 for a description of the legend. Parameter values used for the simulations for these profiles are given in Table S1.

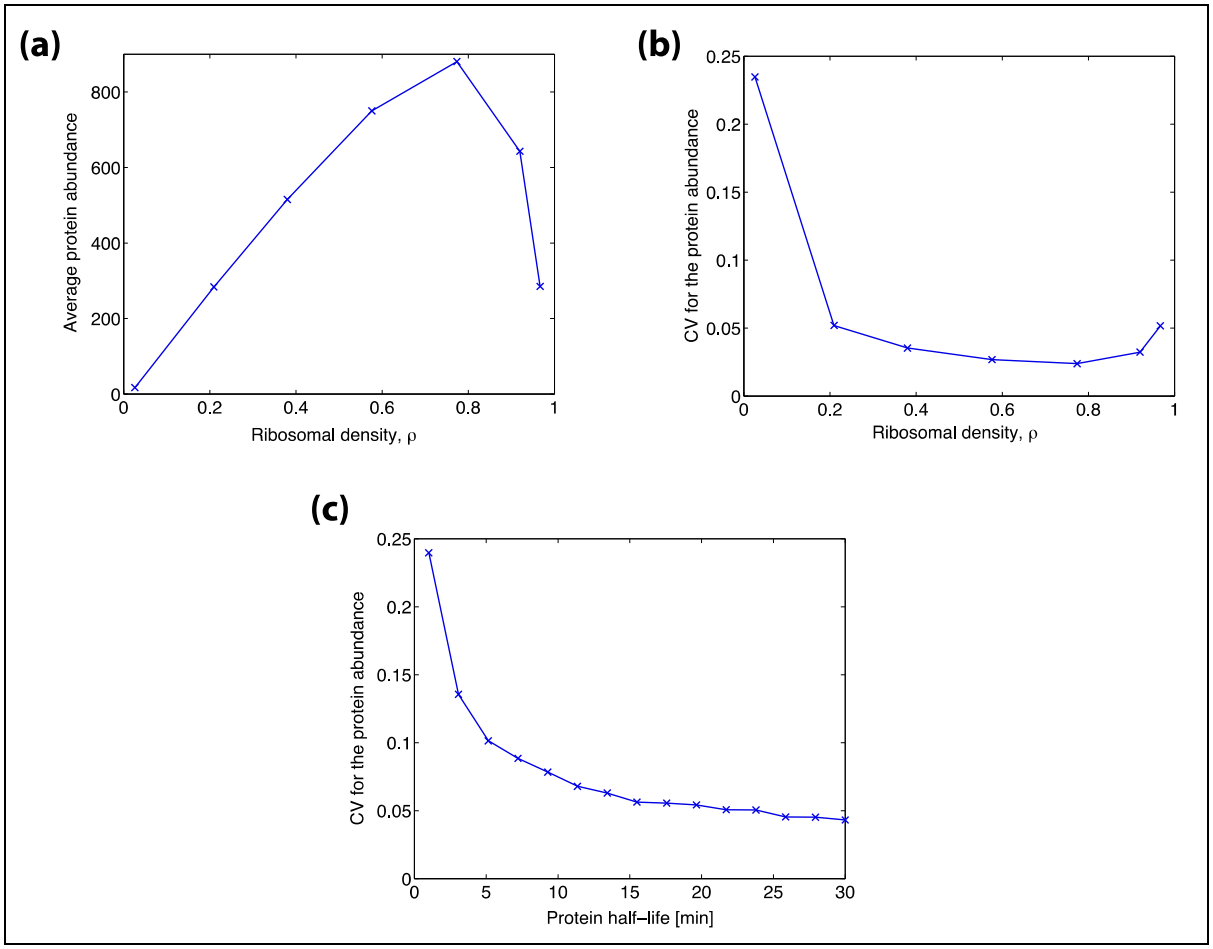


FIGURE S4 Estimating protein abundances and noise on the protein abundance. The simulations are performed as described in Appendix A, and assuming the translation profiles of the genes are the optimal ones (Fig. 3). **(a-b)** Protein abundance (a) and coefficient of variation (b) versus the ribosomal density of the mRNA, for a protein with a half-life of 20 minutes. **(c)** Coefficient of variation for proteins with different half-lives, assuming their mRNA is translated at a ribosomal density of 0.2.

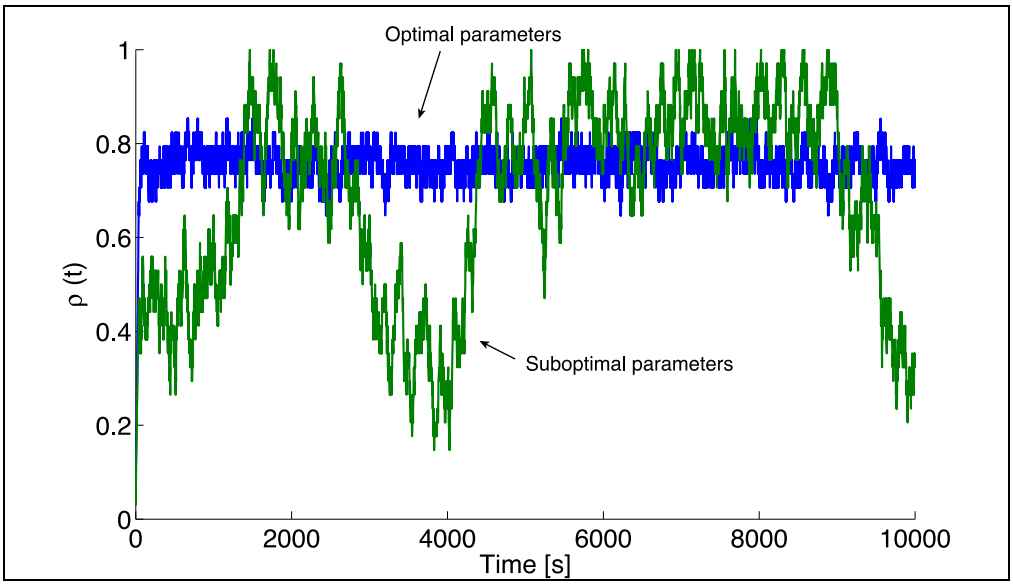


FIGURE S5 Evolution of the instantaneous ribosomal density, with parameters giving maximal synthesis rate (blue curve, $\overline{\rho} = 0.76$) or suboptimal parameters (green curve, $\overline{\rho} = 0.69$).

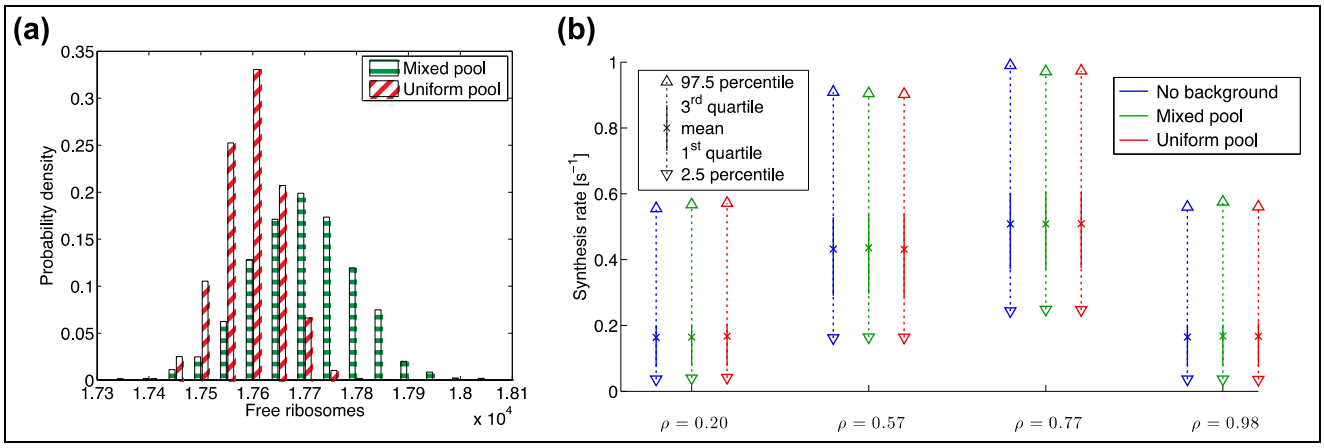


Figure S6 Comparing the results with a background pool of genes or without. **(a)** Distribution of the number of free ribosomes during the evolution of the two background pools of genes. **(b)** Protein synthesis rates when the "marker" gene is isolated or competing with a background pool of genes using two different backgrounds; results are showed for various mean ribosomal densities of the "marker" gene (indicated on the x-axis). Note that for the simulations with the marker gene observed in isolation, a constant number of 17670 free ribosomes was used.

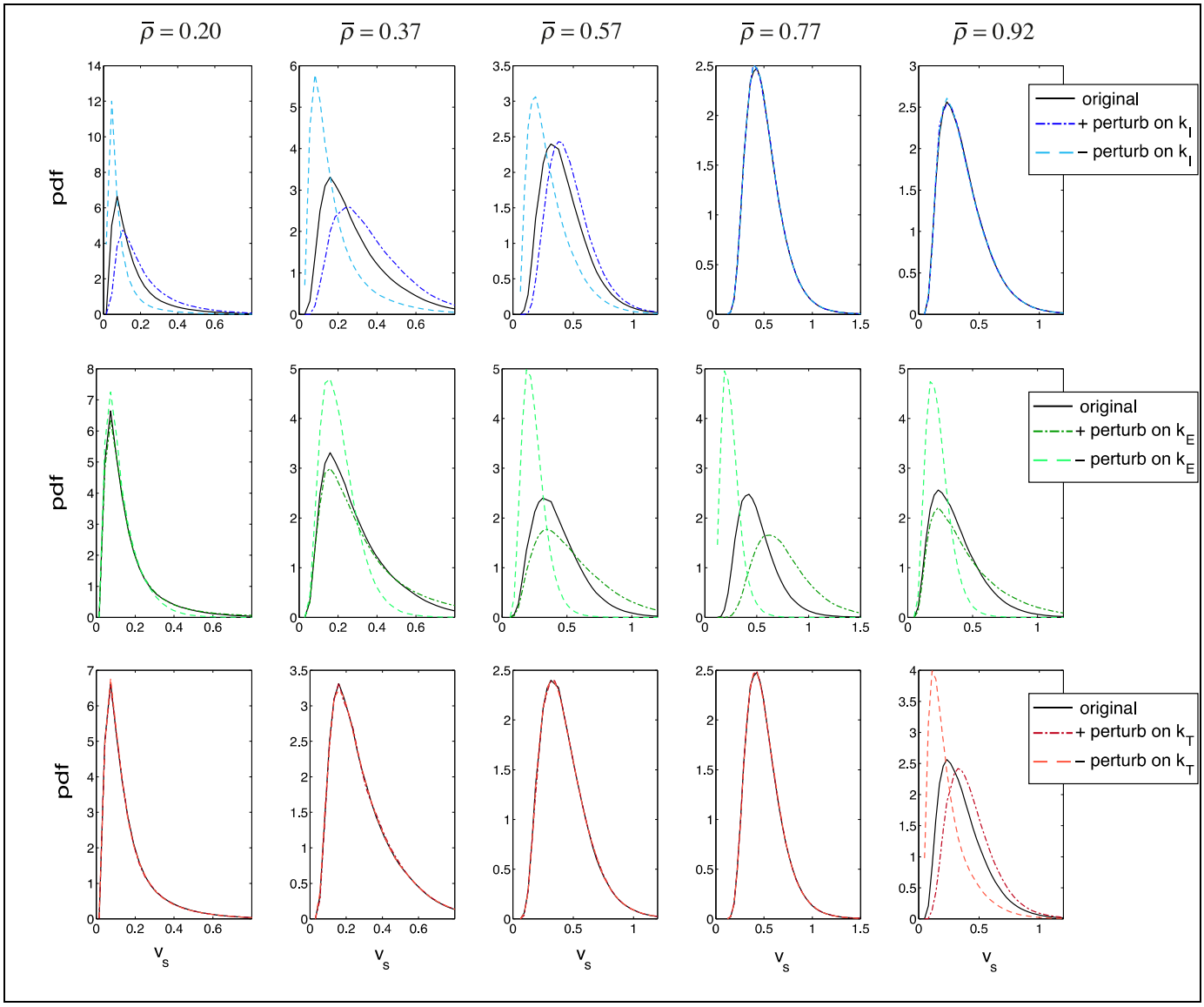


FIGURE S7 Probability distribution of the specific synthesis rates at various densities and resulting distribution after a change of 50% of the given input parameter (given in the legend at the right of each row). The mean densities for the unperturbed cases are given at the top of each column.

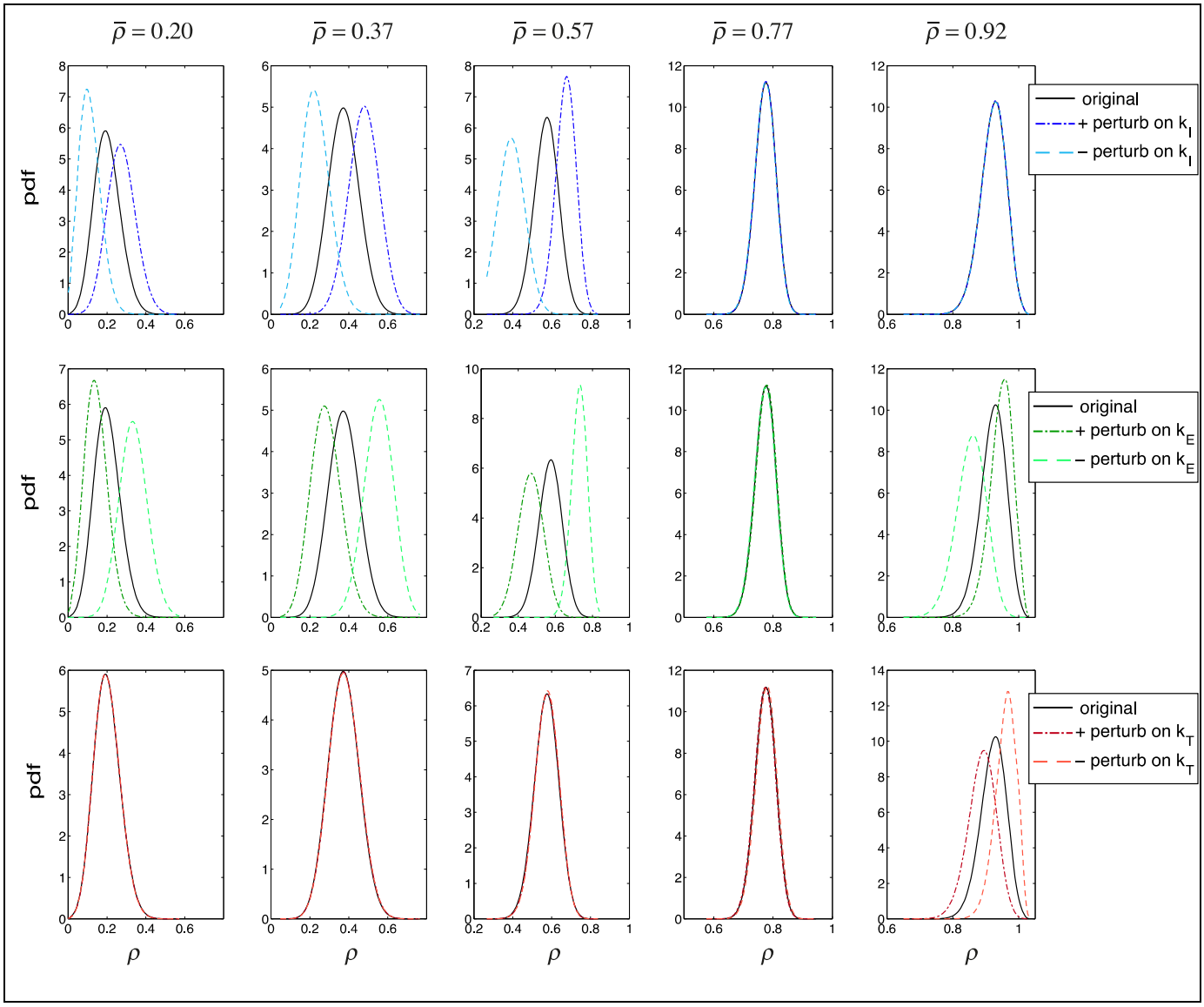
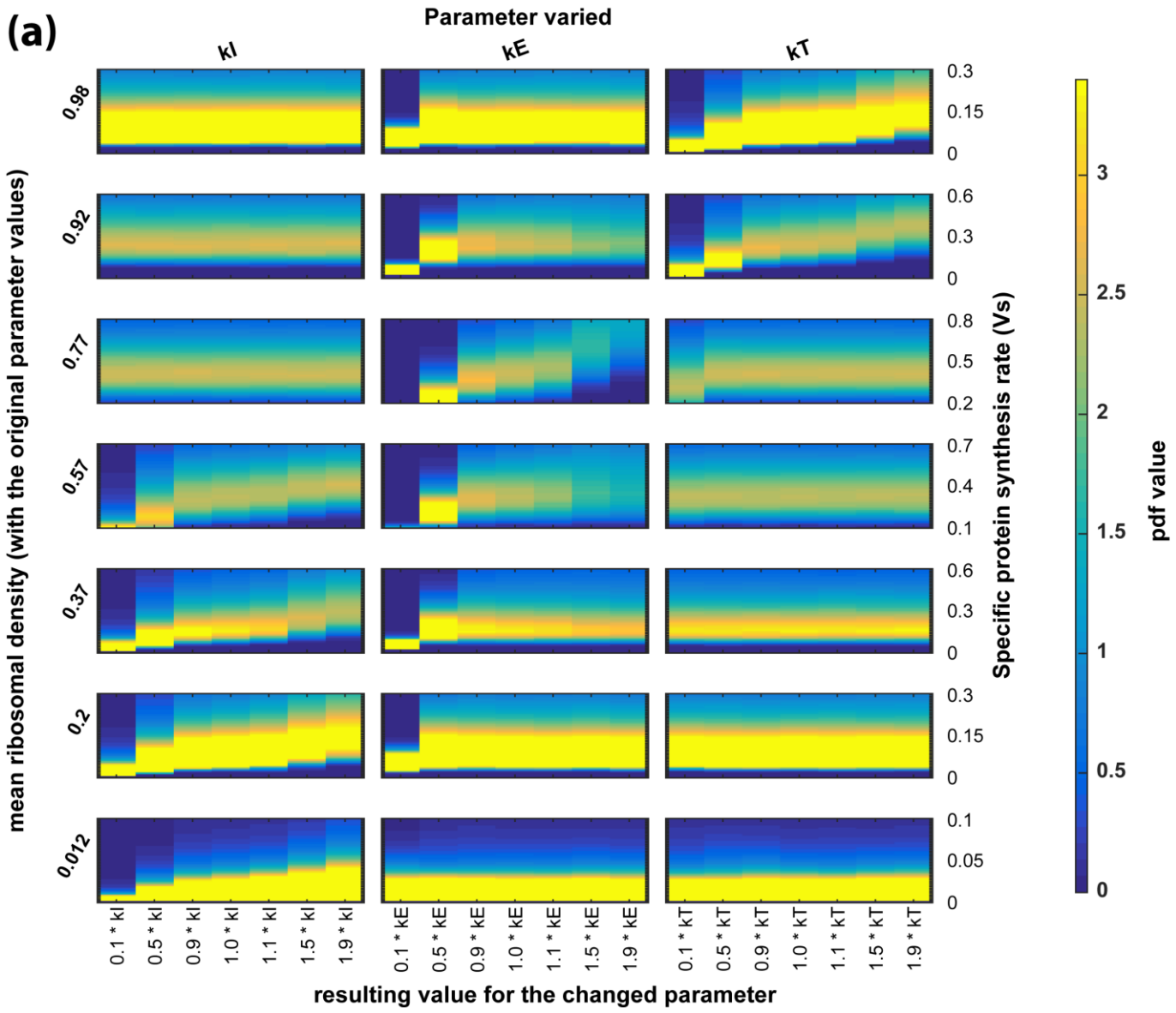


FIGURE S8 Probability distribution of the ribosomal densities at various mean densities and resulting distribution after a change of 50% of the given input parameter (given in the legend at the right of each row). The mean densities for the unperturbed cases are given at the top of each column.


(a)
Parameter varied
kI
kE
kT
0.98
0.92
0.77
0.57
0.37
0.2
0.012
mean ribosomal density (with the original parameter values)
Specific protein synthesis rate (Vs)
pdf value
0.1 * kI
0.5 * kI
0.9 * kI
1.0 * kI
1.1 * kI
1.5 * kI
1.9 * kI
0.1 * kE
0.5 * kE
0.9 * kE
1.0 * kE
1.1 * kE
1.5 * kE
1.9 * kE
0.1 * kT
0.5 * kT
0.9 * kT
1.0 * kT
1.1 * kT
1.5 * kT
1.9 * kT
resulting value for the changed parameter

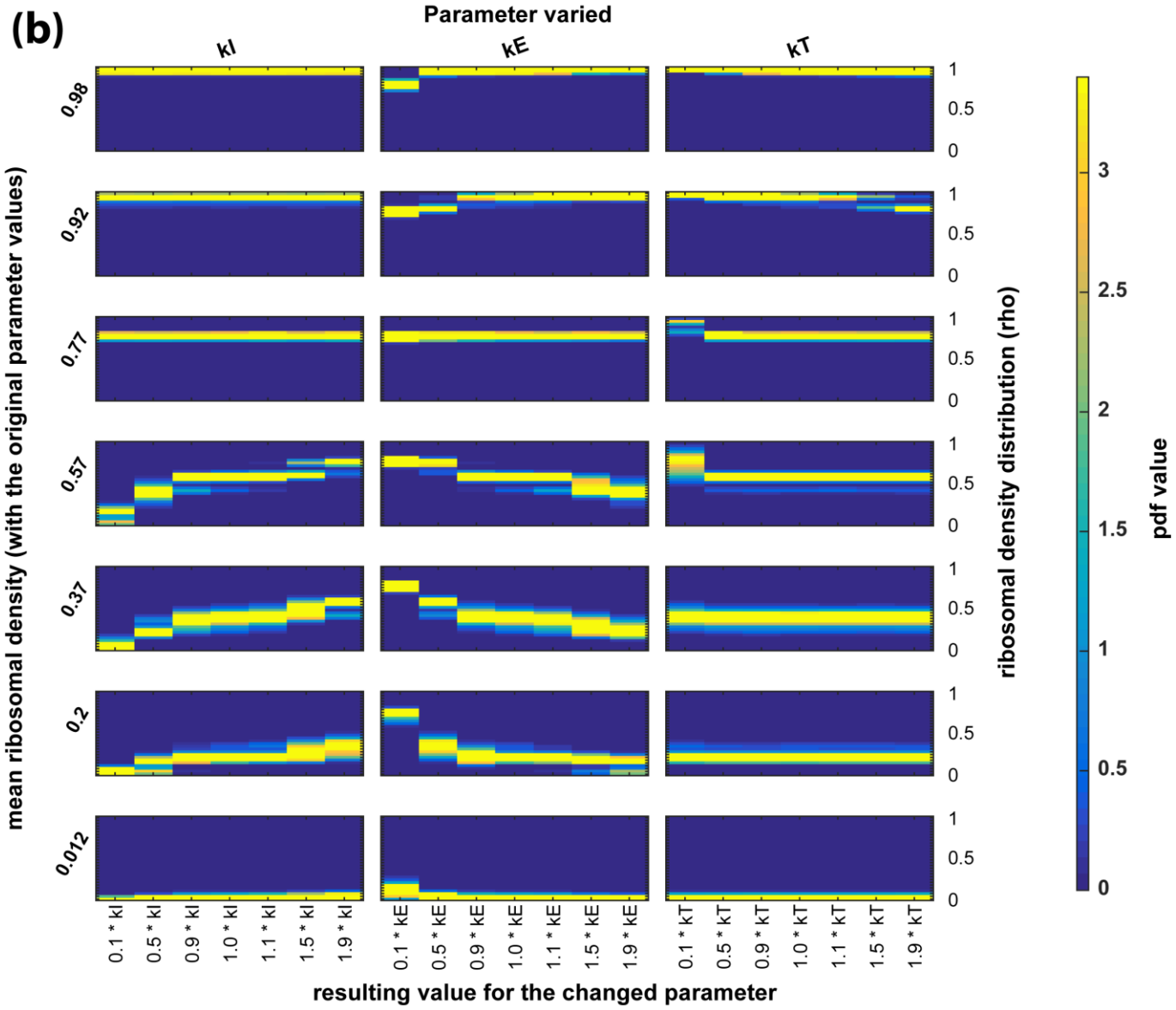


FIGURE S9 Probability density value of the protein synthesis rate (a) and ribosomal densities (b) for various sets of parameters and after different changes on the given input parameters. The input parameters (initiation (1st column), elongation (2nd column) and termination (3rd column) rate constants: kI, kE, kT) were varied one at a time by various amounts (±10, 50 and 90% with respect to the original values), and the resulting *pdf* value for the synthesis rate (a) and ribosomal densities (b) are presented. This was repeated for multiple sets of input parameters that gave rise to different mean ribosomal densities (the various rows of subfigures; the value of ribosomal density given on the left correspond to the mean ribosomal density with the original parameter value sets).

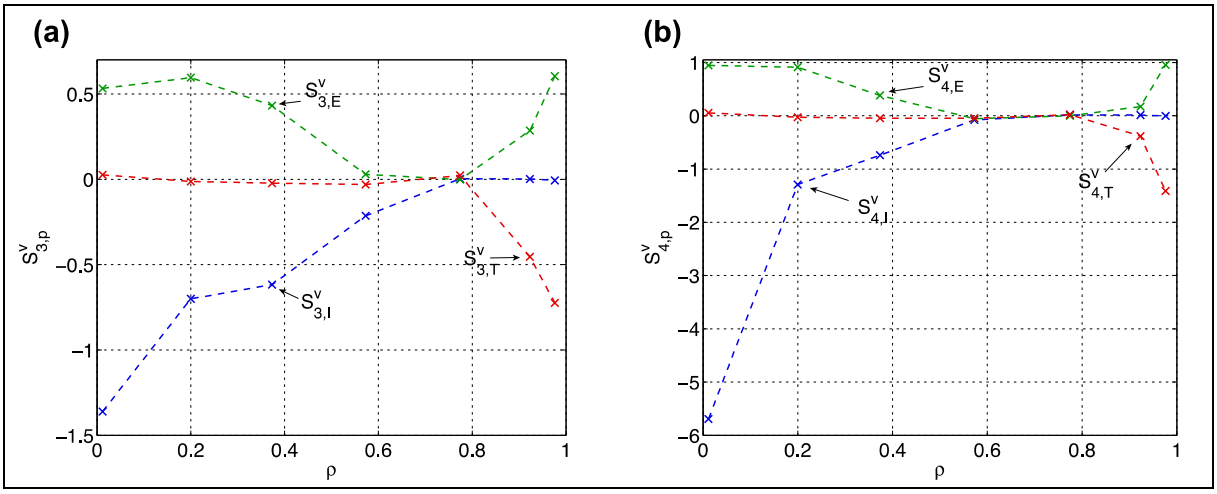


FIGURE S10 Sensitivity of the 3rd moment (a) and 4th moment (b) of the synthesis rate for the initiation rate constant (I), elongation rate constant (E) or termination rate constant (T) in function of the mean ribosomal density.

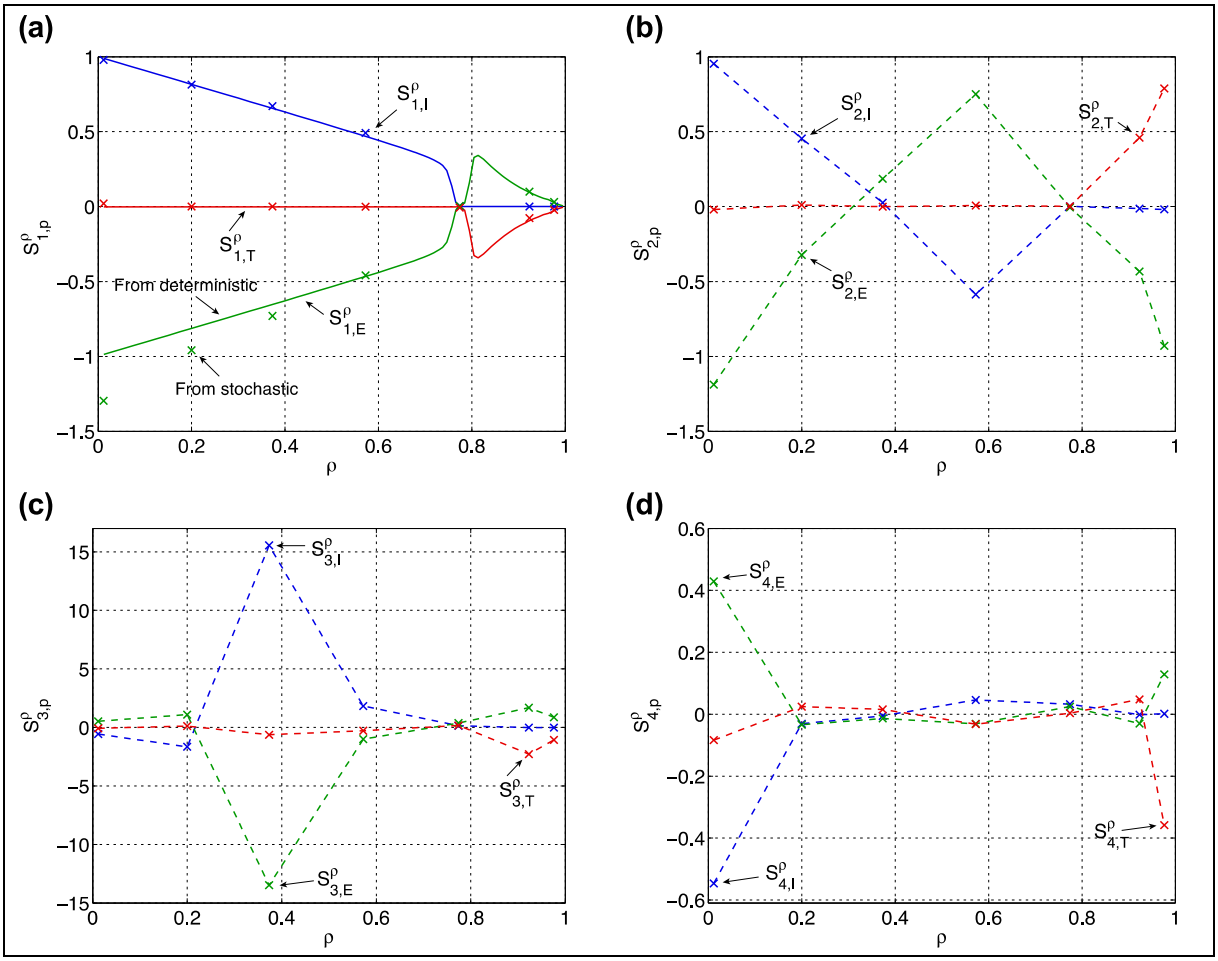


FIGURE S11 Sensitivity of the ribosomal density on the initiation rate constant (*I*), elongation rate constant (*E*) or termination rate constant (*T*) in function of the mean ribosomal density. **(a)** Log-sensitivity for the average ribosomal density (the results from the stochastic simulations are shown as crosses, while the ones computed with the deterministic model are the continuous lines). **(b-d)** Sensitivities of the 2nd, 3rd and 4th moments of the ribosomal density distributions respectively.

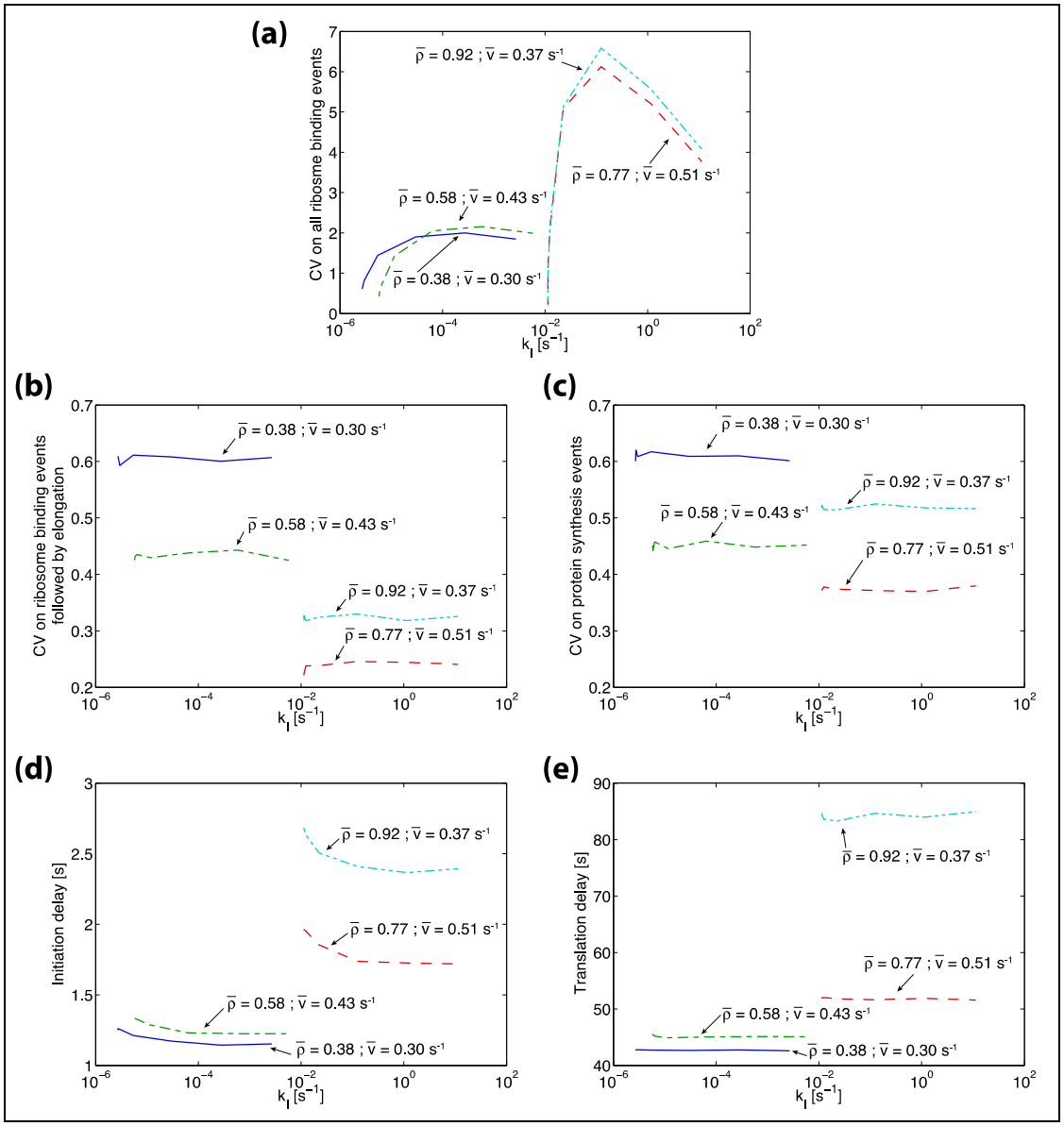


FIGURE S12 Allowing for ribosomes unbinding from initiation site. The rates of translation initiation and reverse-initiation were simultaneously varied in order to keep the same average protein synthesis rates. The corresponding mean ribosomal densities and synthesis rates for 4 different cases are indicated on the figure. The left-most values of k_I for each "line" denotes the minimal k_I value needed to reach the given synthesis rate and ribosomal density (i.e. when $k_{-I} = 0$). **(a)** Coefficient of variation on the rate for all ribosome binding events. **(b)** Coefficient of variation on the rate for the ribosome binding events that are followed by translation elongation (i.e. in (a) all events of initiation are recorded, even those that are followed by the ribosome unbinding from initiation site, while in (b) only the events of initiation that are followed by translation and protein synthesis are recorded). **(c)** Coefficient of variation on the rate of protein synthesis. **(d)** Mean initiation delay, i.e. delay during which the translation initiation site is occupied by a ribosome before this ribosome translated the first L codons of the mRNA, allowing for a new ribosome to bind (the delay reported here only accounts for the ribosomes that perform a full protein translation). **(e)** Mean translation delay, time needed by a ribosome to fully translate the protein, between the translation initiation event and translation termination.

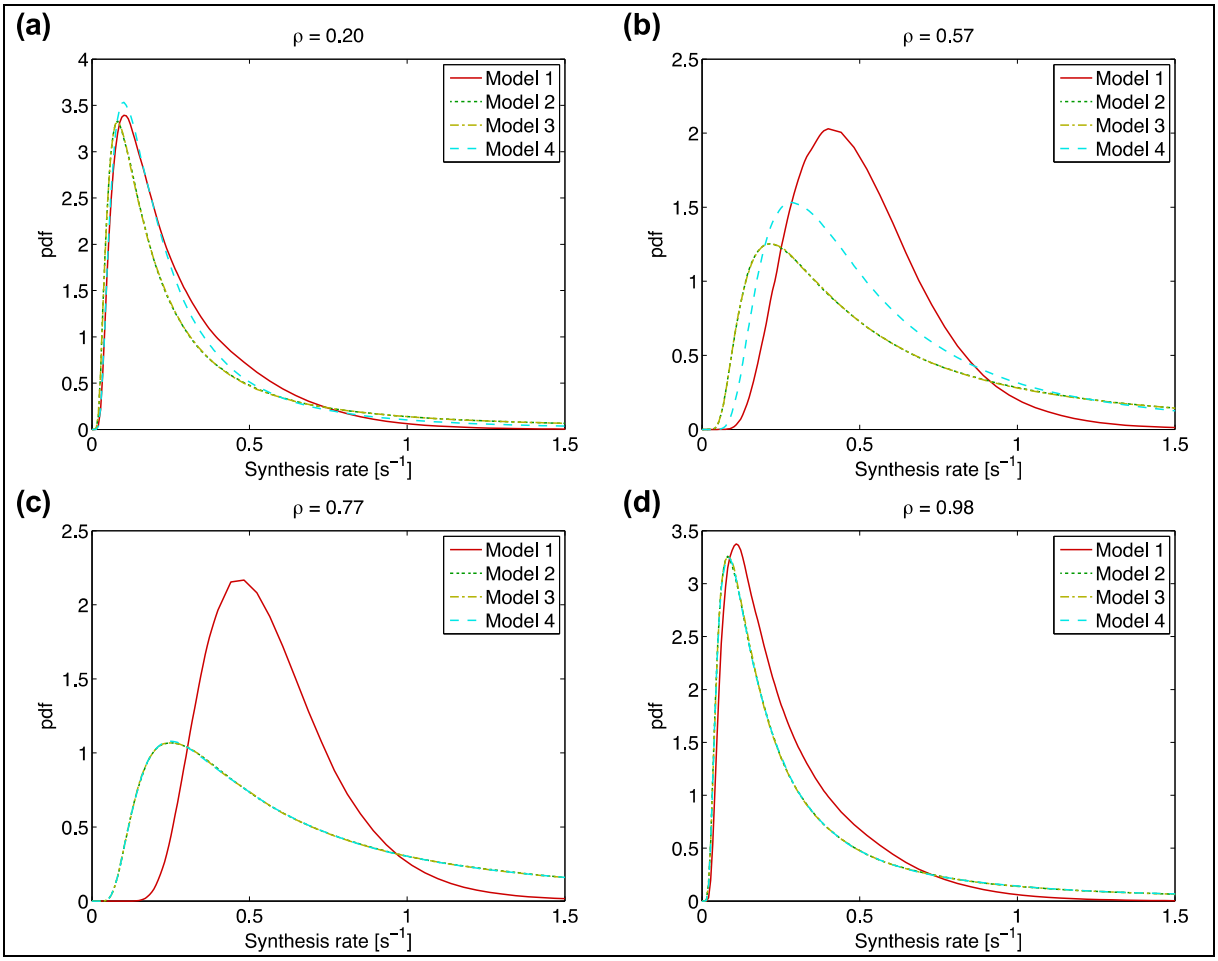


FIGURE S13 Probability density functions of the instantaneous specific synthesis rates for our full model (model 1) and for three simpler models. Showing these distributions at various mean ribosomal densities as indicated in the titles.

Supplemental Tables

Table S1: Parameter values used for the main simulations without a background pool of genes in the case of optimal or suboptimal synthesis profiles. In these simulations the total number of free ribosomes was kept constant. t_{end} describes the time until which the simulations where performed to compute the statistics of protein synthesis. See Table 1 and method section for the parameters definition. As the system is characterized by a single steady state at each parameter set, and the recording starts after the steady state was reached, each state is simulated with a single simulation with a late end-time ($2 \bullet 10^6$ s) which is equivalent to doing for example 1000 repetitions of the simulations during $2 \bullet 10^3$ s.

n [codons]			409
k_E [s^{-1}]			10
R_{free}			17670
$t_{end}[s]$			$2 \cdot 10^6$
Optimal profile	$\rho_{\text{target}} = \mathbf{0.01}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.69 \cdot 10^{-7}$
		k_T [s^{-1}]	10 *
	$\rho_{\text{target}} = \mathbf{0.20}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$1.16 \cdot 10^{-5}$
		k_T [s^{-1}]	10 *
	$\rho_{\text{target}} = \mathbf{0.38}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.70 \cdot 10^{-5}$
		k_T [s^{-1}]	10 *
	$\rho_{\text{target}} = \mathbf{0.58}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.78 \cdot 10^{-5}$
		k_T [s^{-1}]	10 *
	$\rho_{\text{target}} = \mathbf{0.78}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	0.11
		k_T [s^{-1}]	10 *
	$\rho_{\text{target}} = \mathbf{0.93}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	0.11
		k_T [s^{-1}]	0.71
	$\rho_{\text{target}} = \mathbf{0.98}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	0.11
		k_T [s^{-1}]	0.21
Suboptimal profile 1	$\rho_{\text{target}} = \mathbf{0.01}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.69 \cdot 10^{-7}$
		k_T [s^{-1}]	2
	$\rho_{\text{target}} = \mathbf{0.20}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$1.15 \cdot 10^{-5}$
		k_T [s^{-1}]	2

Table S1 continued

	$\rho_{\text{target}} = \mathbf{0.34}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.26 \cdot 10^{-5}$
		k_T [s^{-1}]	2
	$\rho_{\text{target}} = \mathbf{0.58}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.26 \cdot 10^{-5}$
		k_T [s^{-1}]	0.395
	$\rho_{\text{target}} = \mathbf{0.78}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.26 \cdot 10^{-5}$
		k_T [s^{-1}]	0.395
	$\rho_{\text{target}} = \mathbf{0.93}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.26 \cdot 10^{-5}$
		k_T [s^{-1}]	0.393
	$\rho_{\text{target}} = \mathbf{0.98}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$2.26 \cdot 10^{-5}$
		k_T [s^{-1}]	0.162
Suboptimal profile 2	$\rho_{\text{target}} = \mathbf{0.01}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.38 \cdot 10^{-7}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.15}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.61 \cdot 10^{-6}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.36}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.66 \cdot 10^{-6}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.58}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.66 \cdot 10^{-6}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.78}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$5.66 \cdot 10^{-6}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.98}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$1.13 \cdot 10^{-5}$
		k_T [s^{-1}]	0.1
	$\rho_{\text{target}} = \mathbf{0.99}$	k_I [$ribosome^{-1} \cdot s^{-1}$]	$1.13 \cdot 10^{-5}$
		k_T [s^{-1}]	0.05

*: Note that the fact that $k_T = 10 = k_E$ for these cases of the optimal profile is because the termination rate constant is not limiting for all these cases. Any value of $k_T \geq k_E$ give the exact same results for the simulations at these low to medium-high ribosomal densities. Thus we used without loss of generality a value of termination rate constant that was equal to the elongation rate constant for these cases.

Table S2: Parameter values used for the simulations with the *Mixed Pool* background of genes. The background was made up of 10 "gene species" of different copy numbers, length and target ribosomal densities. "Gene 0" is the gene that we studied and we varied its characteristics according to the values from Table S1. Additionally, in these simulations, we used a total number of ribosomes of $R_{tot} = 88352$, an average codon elongation rate $k_E = 10\ s^{-1}$ and simulated the system until $t_{end} = 10^4\ s$.

Gene id	n_i [codons]	# mRNA copies	ρ_{target}	k_I [*]	k_T [s^{-1}]
0	409	1	Values of optimal profile from Table S1		
1	814	2272	0.05	$2.31 \cdot 10^{-6}$	10
2	472	3342	0.15	$8.57 \cdot 10^{-6}$	10
3	290	4885	0.24	$1.52 \cdot 10^{-5}$	10
4	205	1606	0.33	$2.30 \cdot 10^{-5}$	10
5	141	510	0.45	$3.81 \cdot 10^{-5}$	10
6	116	158	0.53	$5.54 \cdot 10^{-5}$	10
7	142	36	0.68	$9.61 \cdot 10^{-5}$	10
8	115	33	0.76	0.11	2.94
9	159	31	0.85	0.11	1.09
10	104	7	0.95	0.11	0.32

[*]: the units of k_I are: [ribosome$^{-1} \cdot s^{-1}$].

Table S3: Parameter values used for the simulations with the *Uniform Pool* background of genes. The gene 0 is the gene that we studied and we varied its characteristics according to the values from Table S1. Additionally, in these simulations, we used $R_{tot} = 88352$, $k_E = 10\ s^{-1}$ and $t_{end} = 10^4\ s$.

Gene id	n_i [codons]	# mRNA copies	ρ_{target}	k_I [*]	k_T [s^{-1}]
0	409	1	Values of optimal profile from Table S1		
1	410	2452	0.82	0.11	1.54

[*]: the units of k_I are: [ribosome$^{-1} \cdot s^{-1}$].