Additional Figures and Complementary Text

Moritz E. Beber[†‡], Patrick Sobetzko[†§], Georgi Muskhelishvili[†] and Marc-Thorsten Hütt[†]

[†]Department of Life Sciences and Chemistry, Jacobs University, Campus Ring 1, 28759 Bremen, Germany
[‡]Bioinformatics Group, Max Planck Institute for Molecular Genetics, Ihnestraße 63–73, 14195 Berlin, Germany
[§]Department of Chromosome Biology, LOEWE-Zentrum für Synthetische Mikrobiologie, Philipps-Universität Marburg, Hans-Meerwein-Straße, 35043 Marburg, Germany

Random Boolean network model

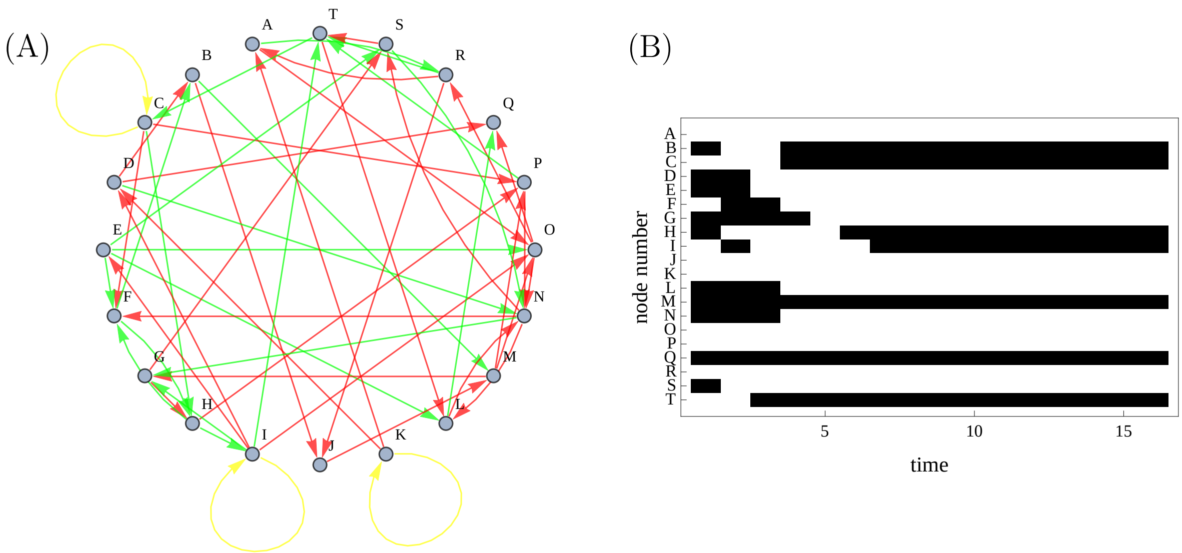

Figure A.1: (A) A small example random Boolean network (RBN). Green links are activating, red ones inhibitory and yellow ones denote self-inhibition. (B) A sample time series of ON/OFF states in the RBN starting from a random state.

Definition of control strengths

Table A.1: Control strengths and their definitions.

Control Strength	Symbol	Equation
Absolute	C_A	Eq. (1)
Functional	C_F	Eq. (2)
Rate	C_R	Eq. (3)
Delay	C_D	Eq. (4)

Figure A.2 shows average control type confidence (CTC) for the four main quantities discussed here. The CTC is a z-score of control strength computed from comparing the original control strength to a random null model. The different control measures are sorted, qualitatively speaking, according to the level of detail incorporated. At the minimal end of the scale, we start with the *absolute control strength* C_A (Section Absolute control strength) which only looks at the difference in expression level between the start and end point of the link under investigation. This is followed by the *functional control strength* C_F (Section Functional control strength) which includes information on the regulatory function of a link (activating or inhibitory) and by the *rate* or *time-difference control strength* C_R (Section Rate control strength) which is based on the change of the gene expression level at the current time point, rather than the expression level itself. Lastly, the *delay control strength* C_D (Section Delay control strength) is shown which is the same as the functional control strength but incorporating a time delay τ_D between start and end node.

The time window for computing the slope entering the rate control strength C_R is one time step. The time delay τ_D entering the delay control strength C_D varies from one to three time steps. We tested, however, that the result does not vary strongly under variation of these parameters. Clearly, at $\tau_D = 0$, $C_D = C_F$. Interestingly, the quantities at intermediate sophistication show the strongest digital control strength. In our subsequent analyses, we use the best performing quantity C_F as a measure for the digital component of the control strength.

It is striking that — after all these processing and normalization steps — our measures of digital control still detect the underlying graph as the generative mechanism (as seen in the positive z-scores for all variants of digital control quantifiers).

Absolute control strength

Measuring analog control strength is the more obvious scenario. Links are based on proximity of genes and we expect neighboring genes to have similar relative expression values due to a high likelihood of being located in the same region of analog control. It seems sensible to regard relative expression levels at the same time point and compute the control strength of a link as follows:

$$C_A(i, j) = 1 - |e_i - e_j| \,. \tag{1}$$

That means, there is a high absolute control strength when relative expression levels e_i and e_j of the neighboring genes i and j are of similar magnitude. For the undirected GPN the distinction of direction has no impact.

Functional control strength

The regulatory function of links in the TRN and thus digital control present much more of a challenge to measure appropriately. The function is mediated by transcription factors which may be liable to any of the following situations: (i) Transcription factors are proteins, that means, they are the result of first transcription, then translation and potentially post-translational modification (PTM). (ii) They may also depend on co-factors for the right conformation. (iii) All these steps may cause a time delay between activation of gene i and a regulatory action at gene j and strongly depend on copy numbers and binding dynamics of the TF.

Obviously, Eq. 1 is a poor approximation of digital control. We can improve it by taking into account the regulatory function of a link. We ignore links that have an unknown or dual role since only a handful of each exist anyway. The following function is applied separately for

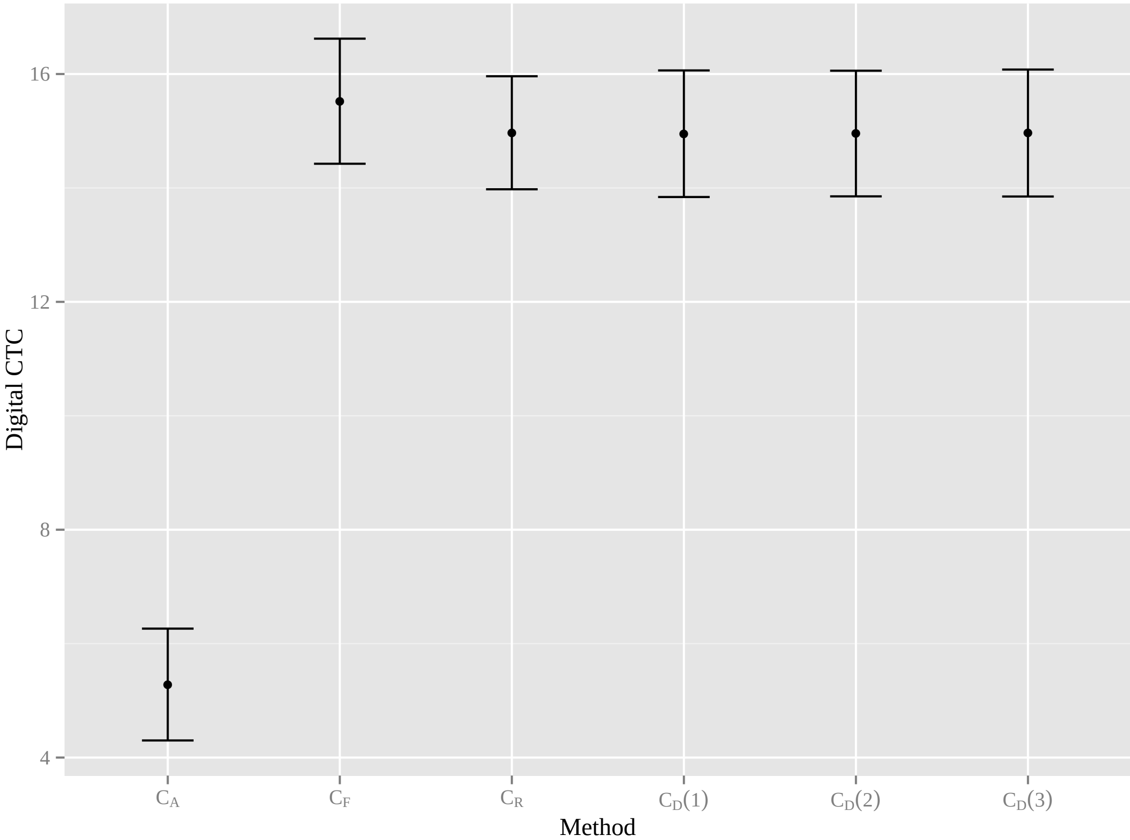

Figure A.2: Average digital control strengths measured by C_A, C_F, C_R and $C_D(1)$, $C_D(2)$ and $C_D(3)$, i.e., for time delays of $\tau_D = 10$, 20 and 30 minutes. The vertical bars depict one standard deviation from the mean in either direction. A discussion of the various control measures can be found in Section Definition of control strengths. The mean functional digital control strength is largest. The RBN results therefore suggest that minimal inclusion of biological information will show the largest signal of digital control.

links A_{ij} that are activating or inhibiting. The matrix A is a variant of the adjacency matrix, which also incorporates the type of the link. The element A_{ij}, denoting the influence of the jth gene on the ith gene, has the following possible entries: 0 (no link), 1 (activating) and -1 (inhibitory).

$$C_F(i,j) = \begin{cases} 1 - |e_i - e_j|\,, & \text{if } A_{ij} = 1 \\ |e_i - e_j|\,, & \text{if } A_{ij} = -1 \end{cases} \tag{2}$$

Rate control strength

A variant of Eq. 2 is to only compare the relative change in expression level between two time points. If the relative change of an activating TF, which is approximated by the relative expression level of its coding gene, is positive, we expect the same for the regulated gene. For an inhibiting link, if the ratio is positive, the regulated gene's ratio should be negative.

$$C_R(i,j) = \begin{cases} 1, & \text{if } A_{ij} = 1 \wedge \delta_j > 0 \wedge \delta_i \geqslant 0 \\ 0, & \text{if } A_{ij} = 1 \wedge \delta_j > 0 \wedge \delta_i < 0 \\ 1, & \text{if } A_{ij} = 1 \wedge \delta_j \leqslant 0 \wedge \delta_i \leqslant 0 \\ 0, & \text{if } A_{ij} = 1 \wedge \delta_j \leqslant 0 \wedge \delta_i > 0 \\ 1, & \text{if } A_{ij} = -1 \wedge \delta_j > 0 \wedge \delta_i \leqslant 0 \\ 0, & \text{if } A_{ij} = -1 \wedge \delta_j > 0 \wedge \delta_i > 0 \\ 1, & \text{if } A_{ij} = -1 \wedge \delta_j \leqslant 0 \wedge \delta_i \geqslant 0 \\ 0, & \text{if } A_{ij} = -1 \wedge \delta_j \leqslant 0 \wedge \delta_i < 0 \end{cases} \tag{3}$$

Delay control strength

Since we not only have to measure the links depending on their function but may also have to take into account delayed action at time points $e_j(t)$ and $e_i(t+\tau_D)$, for example, we can extend the control strength as demonstrated in the following for functional control:

$$C_D(i,j,t) = \begin{cases} 1 - |e_i(t+\tau_D) - e_j(t)|\,, & \text{if } A_{ij} = 1 \\ |e_i(t+\tau_D) - e_j(t)|\,, & \text{if } A_{ij} = -1 \end{cases}. \tag{4}$$

Obviously, such a delay could be introduced for any kind of control strength.

Digital CTC

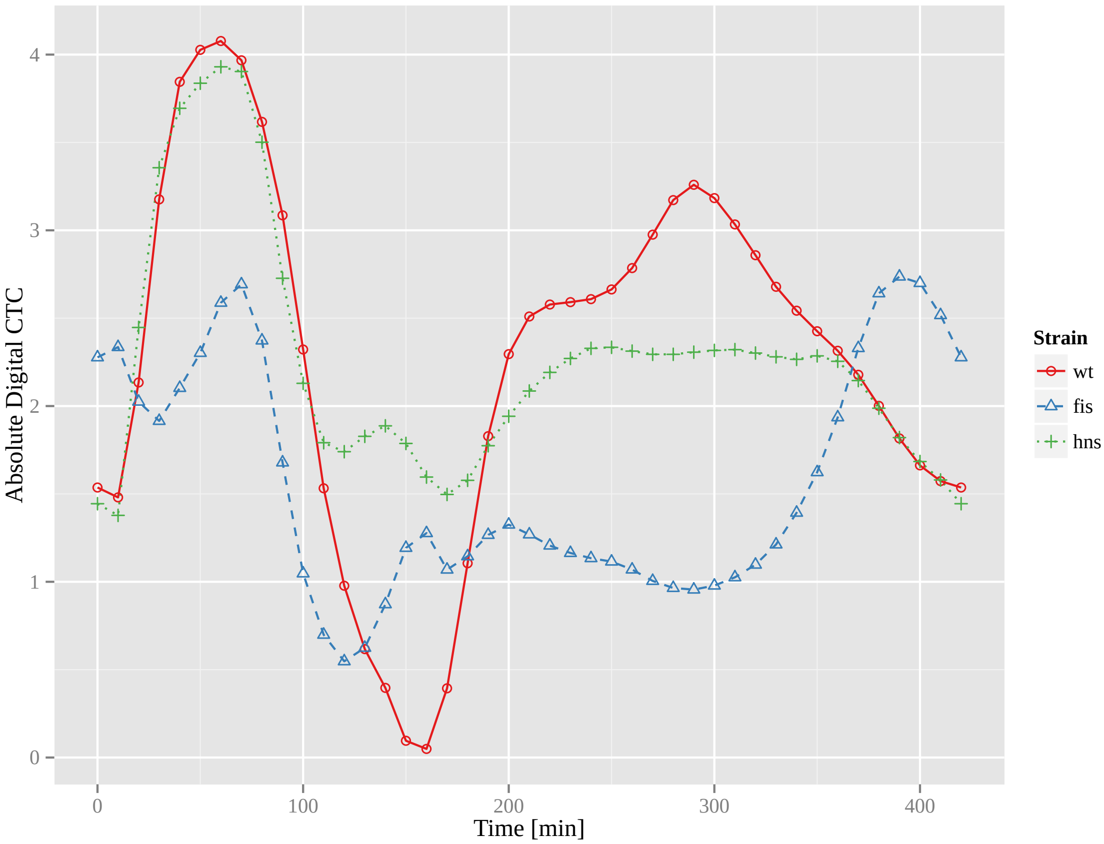

Figure A.3: Digital CTC computed using C_A Eq. 1.

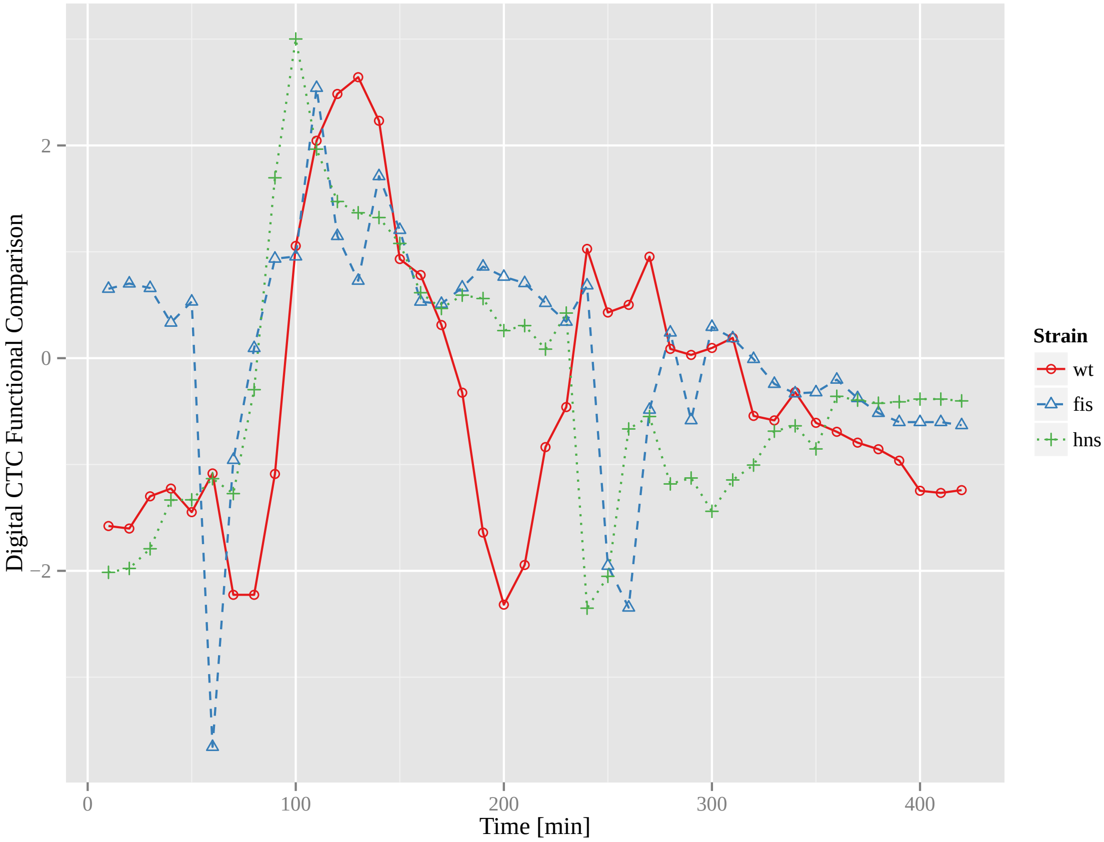

Figure A.4: Digital CTC computed using C_R Eq. 3.

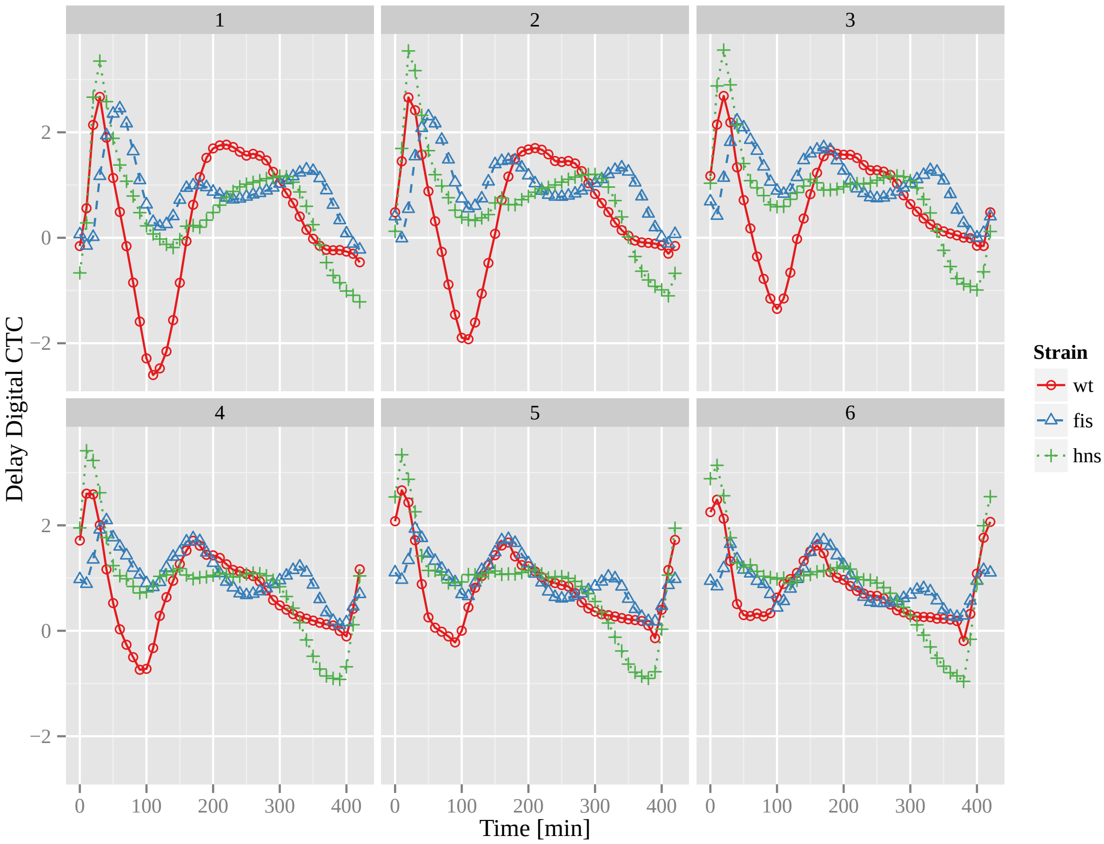

Figure A.5: Digital CTC computed using C_D Eq. 4.