Wireless random-access networks and spectra of random graphs

Sfragara, M.

Author: Sfragara, M.
Title: Wireless random-access networks and spectra of random-graphs
Issue Date: 2020-10-28

PART I

QUEUE-BASED RANDOM-ACCESS PROTOCOLS FOR WIRELESS NETWORKS

CHAPTER 2

Complete bipartite interference graphs

This chapter is based on:
S.C. Borst, F. den Hollander, F.R. Nardi, M. Sfragara. *Transition time asymptotics of queue-based activation protocols for random-access neworks.* Stochastic Processes and Their Applications, 2020.

Abstract

We consider networks where each node represents a server with a queue. An active node deactivates at unit rate. An inactive node activates at a rate that depends on its queue length, provided none of its neighbors is active. For complete bipartite networks, in the limit as the queues become large, we compute the mean transition time between the two states where one half of the network is active and the other half is inactive. We show that the law of the transition time divided by its mean exhibits a trichotomy, depending on the activation rate functions.

§2.1 Introduction and main results

In Section 2.1.1 we describe the setting and the mathematical model of interest in this chapter. In Section 2.1.2 we state our main results. In Section 2.1.3 we offer a brief discussion of these results and give an outline of the remainder of the chapter.

§2.1.1 Setting

We refer to Section 1.1.5 for a general introduction to the mathematical model. In this section we refine it with some extra notions we will need in the chapter.

Consider a complete bipartite graph G: the node set can be partitioned into two nonempty sets U and V such that the bond set is the product of U and V, i.e., two nodes interfere if and only if one belongs to U and the other belongs to V (see Figure 2.1 for an example). Thus, the collection of all independent sets of G consists of all the subsets of U and all the subsets of V.

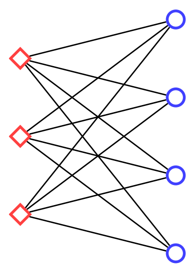

Figure 2.1: A complete bipartite graph with $|U| = 3$ and $|V| = 4$. At time $t = 0$, square-shaped nodes are active and circle-shaped nodes are inactive.

We assume the activation rates to satisfy Definition 1.1.4. Moreover, we focus on the following.

Definition 2.1.1 (Assumption on the activation rates).
We assume *polynomial activation functions* for nodes in U of the form

$$g_U(x) \sim Bx^\beta, \qquad x \to \infty, \tag{2.1}$$

with $B, \beta \in (0, \infty)$. We will discuss more general functions g_U in Remark 2.4.1. We do not require any further assumption on the functions g_V: it turns out that the asymptotic distribution of the transition time is independent of g_V.

Next, we define our two main objects of interest.

Definition 2.1.2 (Pre-transition and transition time).
The *pre-transition time* τ_G is defined as the first time a node in V activates starting from u, i.e.,

$$\tau_G = \inf\{t > 0\colon\ X_i(t) = 1\ \exists\, i \in V,\ X(0) = u\}. \tag{2.2}$$

The *transition time* $\mathcal{T}_G$ is defined as the first time v is reached starting from u, i.e.,

$$\mathcal{T}_G = \inf\big\{t \geq 0\colon\ X_i(t) = 0\ \forall\, i \in U,\ X_i(t) = 1\ \forall\, i \in V, X(0) = u\big\}. \tag{2.3}$$

The pre-transition time plays an important role in our analysis of the transition time, because the evolution of the network is simpler on the interval $[0, \tau_G]$ than on the interval $[\tau_G, \mathcal{T}_G]$. However, we will see that $\mathcal{T}_G - \tau_G \ll \tau_G$ when the initial queue lengths are large, so that both times have the same asymptotic scaling behavior. See Figure 2.2 for a representation of the pre-transition state.

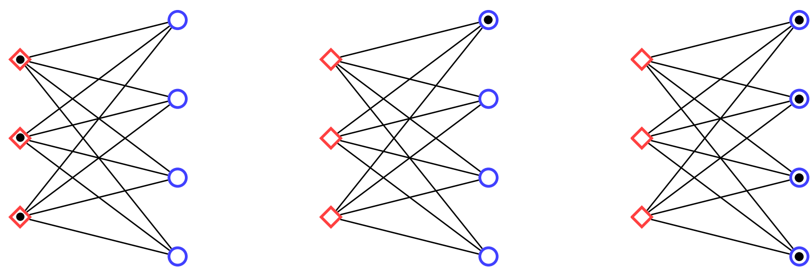

Figure 2.2: Left*: initial state* u*.* Center*: pre-transition state.* Right*: final state* v*.*

We study the transition starting from u and we set the initial queue sizes $Q_i(0)$ to be large for all $i \in U \sqcup V$. Hence, initially all the nodes in U are active virtually all the time, preventing any of the nodes in V to activate. Consequently, the queue sizes of the nodes in U will tend to decrease at rate $c - \rho_U > 0$, while the queue sizes of the nodes in V will tend to increase at rate $\rho_V > 0$. While the packet arrivals and activity periods are governed by random processes, the trajectories of the queue sizes will be roughly linear when viewed on the long time scales of interest.

As mentioned in Section 1.1.4, we focus on queue-based random-access protocols where the activation rates are functions of the queue lengths at the various nodes. We call these protocols internal models and in particular we study the internal model with activation rates as described in (1.5). Since we assume identical initial queue sizes within the sets U and V, the asymptotic distribution of the transition time in the internal model should be close to that in the external model described in (1.6) when we choose

$$h_U(t) = g_U\big(Q_U(0) - (c - \rho_U)t\big), \qquad h_V(t) = g_V\big(Q_V(0) + \rho_V t\big), \tag{2.4}$$

with $Q_U(0) = \gamma_U r$ and $Q_V(0) = \gamma_V r$. Next, we formalize our four main models of interest.

Definition 2.1.3 (Models).
Let $\delta > 0$.

- In the *internal model* the deactivation Poisson clocks tick at rate 1, while the activation Poisson clocks tick at rate

$$r_i^{\text{int}}(t) = \begin{cases} g_U(Q_i(t)), & i \in U, \\ g_V(Q_i(t)), & i \in V, \end{cases} \qquad t \geq 0. \tag{2.5}$$

- In the *external model* the deactivation Poisson clocks tick at rate 1, while the activation Poisson clocks tick at rate

$$r_i^{\mathrm{ext}}(t) = \begin{cases} g_U(\gamma_U r - (c-\rho_U)t), & i \in U, \\ g_V(\gamma_V r + \rho_V t), & i \in V, \end{cases} \qquad t \geq 0. \tag{2.6}$$

- In the *lower external model* the deactivation Poisson clocks tick at rate 1, while the activation Poisson clocks tick at rate

$$r_i^{\mathrm{low}}(t) = \begin{cases} g_U(\gamma_U r - (c-\rho_U)t - \delta r), & i \in U, \\ g_V(\gamma_V r + \rho_V t + \delta r), & i \in V, \end{cases} \qquad t \geq 0. \tag{2.7}$$

- In the *upper external model* the deactivation Poisson clocks tick at rate 1, while the activation Poisson clocks tick at rate

$$r_i^{\mathrm{upp}}(t) = \begin{cases} g_U(\gamma_U r - (c-\rho_U)t + 2\delta r), & i \in U, \\ g_V(\gamma_V r + \rho_V t - \delta r), & i \in V, \end{cases} \qquad t \geq 0. \tag{2.8}$$

Note that in the three external models the activation rates depend on time via certain fixed parameters, while in the internal model they depend on time via the actual queue lengths at the nodes. In the lower external model the activation rates in U tend to be less aggressive than in the internal model (i.e., the activation clocks tick less frequently), while the activation rates in V tend to be more aggressive. In the upper external model the reverse is true: the activation rates in U are more aggressive and the activation rates in V are less aggressive. For simplicity, when considering the external model we sometimes write $r_U(t)$ and $r_V(t)$ for the activation rates at time t of nodes in U and nodes in V, respectively. We will see that the upper external model is actually defined only for $t \in [0, T_U]$ with $T_U = \frac{\gamma_U}{c-\rho_U} r$ (see Section 2.2 for details). However, with high probability as $r \to \infty$, the transition occurs before time T_U.

§2.1.2 Main theorems

The main goal of the chapter is to compare the transition time of the internal model with that of the external model. Through a large deviation analysis of the queue length process at each of the nodes, we define a notion of *good behavior* that allows us to define perturbed models with externally driven activation rates that sandwich the queue lengths of the internal model and its transition time. We show with the help of coupling that, with high probability as $r \to \infty$, the asymptotic behavior of the mean transition time for the internal model is the same as for the external model.

The metastable behavior and the transition time $\mathcal{T}_G$ of a network in which the activation rates are time-dependent in a deterministic way was characterized in [14], with the help of the metastability analysis for hard-core interaction models developed in [59]. For $s \geq 0$, let

$$\nu(s) = \frac{1}{\mathbb{E}_u[\mathcal{T}_G](s)} \tag{2.9}$$

be the inverse mean transition time of the time-homogeneous model where we freeze the the activation rates r_U and r_V at time s, i.e., we consider the model with constant activation rates

$$r_i^{\text{ext}}(t) = \begin{cases} r_U(s), & i \in U, \\ r_V(s), & i \in V, \end{cases} \qquad t \geq 0. \tag{2.10}$$

Then, for any time scale $M = M(r)$ and any threshold $x \in [0, \infty)$,

$$\lim_{r\to\infty} \mathbb{P}_u\left(\frac{\mathcal{T}_G}{M} > x\right) = \begin{cases} 0, & \text{if } M\nu(Mx) \succ 1, \\ e^{-\int_0^x M\nu(Ms)ds}, & \text{if } M\nu(Mx) \asymp 1, \\ 1, & \text{if } M\nu(Mx) \prec 1. \end{cases} \tag{2.11}$$

(Here, as $r \to \infty$, $a \succ b$ means $b = o(a)$, $a \prec b$ means $a = o(b)$, while $a \asymp b$ means $a = \Theta(b)$.) If we let M_c be the unique solution of the equation

$$M\nu(M) = 1, \tag{2.12}$$

then the transition occurs on the time scale M_c, in the sense that $\mathbb{P}_u(\mathcal{T}_G > t) \approx 1$ for $t \prec M_c$ and $\mathbb{P}_u(\mathcal{T}_G > t) \approx 0$ for $t \succ M_c$. On the critical time scale M_c, the transition time follows an exponential distribution with time-varying rate. It was proven in [59] that, for a complete bipartite graph and $s \in [0, \infty)$,

$$\mathbb{E}_u[\mathcal{T}_G](s) = \frac{1}{|U|}\, r_U(s)^{|U|-1}\, [1 + o(1)], \qquad r \to \infty. \tag{2.13}$$

The following two theorems will be proven in Sections 2.4.1–2.4.2 with the help of (2.9)–(2.13).

Theorem 2.1.4 (Critical time scale in the external model).
The time scale on which the transition occurs is given by

$$M_c = F_c\, r^{1\wedge\beta(|U|-1)}\, [1 + o(1)], \qquad r \to \infty, \tag{2.14}$$

with

$$F_c = \begin{cases} \frac{\gamma_U^{\beta(|U|-1)}}{|U|B^{-(|U|-1)}}, & \text{if } \beta \in (0, \frac{1}{|U|-1}), \\ \frac{\gamma_U}{|U|B^{-(|U|-1)}+(c-\rho_U)}, & \text{if } \beta = \frac{1}{|U|-1}, \\ \frac{\gamma_U}{c-\rho_U}, & \text{if } \beta = (\frac{1}{|U|-1}, \infty). \end{cases} \tag{2.15}$$

Theorem 2.1.5 (Transition time in the external model).
The transition time in the external model satisfies

$$\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}] = F_c\, r^{1\wedge\beta(|U|-1)}\, [1 + o(1)], \qquad r \to \infty. \tag{2.16}$$

with F_c as in (2.15)*, and*

$$\lim_{r\to\infty} \mathbb{P}_u\left(\frac{\mathcal{T}_G^{\text{ext}}}{\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}]} > x\right) = \mathcal{P}(x), \qquad x \in [0, \infty), \tag{2.17}$$

with

$$\mathcal{P}(x) = \begin{cases} e^{-x}, & \text{if } \beta \in (0, \frac{1}{|U|-1}),\ x \in [0, \infty), \\ (1 - Cx)^{\frac{1-C}{C}}, & \text{if } \beta = \frac{1}{|U|-1},\ x \in [0, \frac{1}{C}), \\ 0, & \text{if } \beta = \frac{1}{|U|-1},\ x \in [\frac{1}{C}, \infty), \\ 1, & \text{if } \beta \in (\frac{1}{|U|-1}, \infty),\ x \in [0, 1), \\ 0, & \text{if } \beta \in (\frac{1}{|U|-1}, \infty),\ x \in [1, \infty), \end{cases} \tag{2.18}$$

and $C = \frac{F_c(c-\rho_U)}{\gamma_U} \in (0,1)$.

In other words, the mean transition time scales like M_c, while the law of the transition time divided by its mean is exponential, truncated polynomial or deterministic (see Figure 2.3). We distinguish between these three regimes of behavior and refer to them as *subcritical regime*, *critical regime* and *supercritical regime*, respectively. The deterministic behavior observed in the supercritical regime is also known in the literature as *cut-off*.

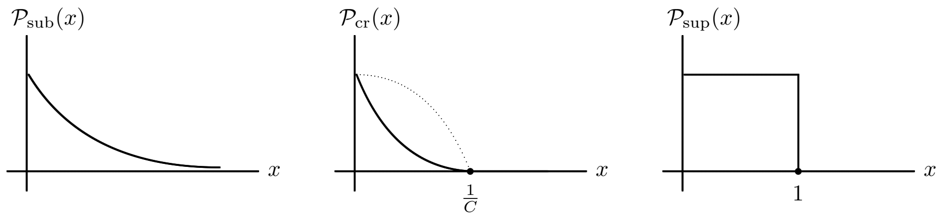

Figure 2.3: Trichotomy for $x \mapsto \mathcal{P}(x)$: $\beta \in (0, \frac{1}{|U|-1}]$, subcritical regime (left); $\beta = \frac{1}{|U|-1}$, critical regime (middle); $\beta \in (\frac{1}{|U|-1}, \infty)$, supercritical regime (right). The curve in the middle is convex when $C \in (0,1/2)$ and concave when $C \in (1/2,1)$. The curve on the right is the limit of the curve in the middle as $C \to 1$.

As shown in Remark 2.4.1, we can even include the case $\beta = 0$, and get that if $g_U(x) = \hat{\mathcal{L}}(x)$ with $\lim_{x\to\infty} \hat{\mathcal{L}}(x) = \infty$, then

$$\mathbb{E}_u[\mathcal{T}_G^{\rm ext}] = M_c\,[1+o(1)], \qquad M_c = \frac{1}{|U|}\hat{\mathcal{L}}(\gamma_U r)^{|U|-1}\,[1+o(1)], \qquad r\to\infty, \quad (2.19)$$

and $\mathcal{P}(x) = e^{-x}$, $x \in [0,\infty)$. Similar properties hold for the lower and the upper external model, with perturbed $F_{c,\delta}^{\rm low}$ and $F_{c,\delta}^{\rm upp}$ satisfying

$$\lim_{\delta\to 0} F_{c,\delta}^{\rm low} = \lim_{\delta\to 0} F_{c,\delta}^{\rm upp} = F_c. \quad (2.20)$$

The main result of the chapter is the following sandwich of $\mathcal{T}_G^{\rm int}$ between $\mathcal{T}_G^{\rm low}$ and $\mathcal{T}_G^{\rm upp}$, for which we already know the asymptotic behavior. Because of this sandwich we can deduce the asymptotics of the transition time in the internal model.

Theorem 2.1.6 (Transition time in the internal model).
For $\delta > 0$ small enough, there exists a coupling such that

$$\lim_{r\to\infty} \hat{\mathbb{P}}_u\big(\mathcal{T}_G^{\rm low} \le \mathcal{T}_G^{\rm int} \le \mathcal{T}_G^{\rm upp}\big) = 1, \quad (2.21)$$

where $\hat{\mathbb{P}}_u$ is the joint law induced by the coupling, with all three models starting from u. Consequently, the transition time in the internal model satisfies

$$\mathbb{E}_u[\mathcal{T}_G^{\rm int}] = F_c\, r^{1\wedge\beta(|U|-1)}\,[1+o(1)], \qquad r\to\infty, \quad (2.22)$$

with F_c as in (2.15)*, and*

$$\lim_{r\to\infty} \mathbb{P}_u\left(\frac{\mathcal{T}_G^{\mathrm{int}}}{\mathbb{E}_u[\mathcal{T}_G^{\mathrm{int}}]} > x\right) = \mathcal{P}(x), \qquad x \in [0,\infty). \tag{2.23}$$

with $\mathcal{P}(x)$ as in (2.18)*.*

§2.1.3 Discussion and outline

Theorems. Theorem 2.1.5 gives the leading-order asymptotics of the transition time in the external model, including the lower and the upper external model. Theorem 2.1.6 is the main result of the chapter and provides the leading-order asymptotics of the transition time in the internal model, via the coupling in (2.21) and the continuity property in (2.20). Equations (2.15)–(2.16) identify the scaling of the transition time in terms of the model parameters. The trichotomy between $\beta \in (0, \frac{1}{|U|-1})$, $\beta = \frac{1}{|U|-1}$ and $\beta \in (\frac{1}{|U|-1}, \infty)$ is particularly interesting, and leads to different limit laws for the transition time on the scale of its mean.

Interpretation of the trichotomy. In order to interpret the above trichotomy, observe first of all that the activation rates of each of the nodes in U remain of order r^β almost all the way up T_U. Specifically, in the absence of the nodes in V, by time yT_U, $y \in [0,1)$, the queue lengths of the nodes in U have decreased by roughly a fraction y, and their activation rates are approximately $B(1-y)^\beta r^\beta$. Hence the fraction of joint inactivity time of the nodes in U is of order $(1/r^\beta)^{|U|} = r^{-\beta|U|}$. Since the tike it takes to leave the joint inactivity state is of order $r^{-\beta}$, all nodes in U become simultaneously inactive for the first time after a period of order $r^{-\beta}/r^{-\beta|U|} = r^{\beta(|U|-1)}$, which is $o(r)$ in the subcritical regime when $\beta < \frac{1}{|U|-1}$. When the nodes in V are actually present, with high probability as $r \to \infty$, they all activate quickly and the transition occurs almost immediately (see Section 2.4.3). Note that the queue lengths of the nodes in U have only decreased by an amount of order $r^{\beta(|U|-1)} = o(r)$, and hence are still of order r. In contrast, in the critical regime when $\beta = \frac{1}{|U|-1}$, the probability that all nodes in U become simultaneously inactive before time yT_U is approximately $\pi(y)$ with $\pi(y) = 1-(1-y)^{(1-C)/C}$, $y \in [0,1)$ (see (2.18)). Again, with high probability as $r \to \infty$, all the nodes in V activate quickly and the transition occurs almost immediately. Note that the queue lengths in the nodes in U have then dropped by a non-negligible fraction, but are still of order r. A potential scenario is that the nodes in U do not all become simultaneously inactive until their activation rates have become of a smaller order than r^β, due to the queue lengths no longer being of order r just before time T_U. However, the fact that $\pi(y) \to 1$ as $y \to 1$ implies that this scenario has negligible probability in the limit. In contrast, this scenario does occur in the supercritical regime when $\beta > \frac{1}{|U|-1}$, implying that the crossover occurs in a narrow window around T_U (see Sections 2.4.1–2.4.2 for details). We will see that this window has size $O(r^{1/\beta(|U|-1)}) = o(r)$. In particular, the window gets narrower as the activation rates of nodes in U increase.

Proofs. We look at a single-node queue length process $t \mapsto Q(t)$ and prove that with high probability it follows a path that lies in a narrow tube around its mean path (see

Figure 2.4). We study separately the input process $t \mapsto Q^+(t)$ and the output process $t \mapsto Q^-(t)$: we use Mogulskii's theorem (a pathwise large deviation principle) for the first, and Cramér's theorem (a pointwise large deviation principle) for the second. We derive upper and lower bounds for the queue length process and we use these bounds to construct two couplings that allow us to compare the different models.

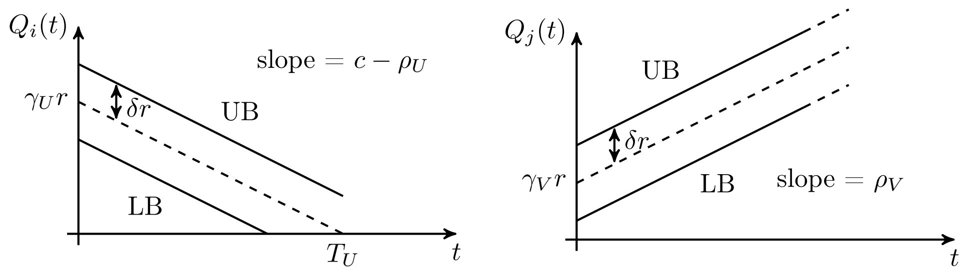

Figure 2.4: Sketches of the tubes around the mean of the queue length processes, respectively, for a node $i \in U$ and a node $j \in V$.

Dependent packet arrivals. Our large deviation estimates are so sharp that we can actually allow the Poisson processes of packet arrivals at the different nodes to be dependent. Indeed, as long at the marginal processes are Poisson, our large deviation estimates are valid at every single node, and since the network is finite a simple union bound shows that they are also valid for all nodes simultaneously, at the expense of a negligible factor that is proportional to the number of nodes. For modeling purposes independent arrivals are natural, but it is interesting to allow for dependent arrivals when we want to study activation protocols that are more involved.

Open problems. If we want to understand how small the term $o(1)$ in (2.22) actually is, then we need to derive sharper estimates in the coupling. One possibility would be to study moderate deviations for the queue length processes and to look at shrinking tubes. We do not pursue such refinements here. Our main focus for the future will be to extend the model to more complicated settings, where the activation rate at node i depends also on the queue length at the neighboring nodes of i. We want to be able to compare models with (externally driven) time-dependent activation rates and models with (internally driven) queue-dependent activation rates, and show again that their metastable behavior is similar. We also want to move away from the complete bipartite interference graph and consider more general graphs that capture more realistic wireless networks.

Other models. There are other ways to define an internal model. We mention a few examples.

(i) A simple variant of our model is obtained by fixing the activation rates, but letting the rate at time t of the Poisson deactivation clock of node i depend on the reciprocal of the queue length at time t, i.e., $1/g_i(Q_i(t))$ for some $g_i \in \mathcal{G}$. This can be equivalently seen as a unit-rate Poisson deactivation clock, where node i either deactivates with a probability reciprocal to $g_i(Q_i(t))$, or starts a second activity period. Nodes with a large queue length are more likely to remain active for a long time before deactivating, while nodes with a short queue length have extremely short activity periods. If at time t the activation clock of an inactive node with $Q_i(t) = 0$ ticks, then the node does not activate. On the other hand, if during an activity period the queue length of an active node hits zero, then the node deactivates independently of its deactivation rate. For fixed activation and deactivation rates, this model and our internal model with $r_i^{\mathrm{int}}(t) = g_i(Q_i(t))$ for each node i are equivalent up to a time scaling factor. In particular, they have similar stationary distributions.

(ii) An alternative approach is to use a discrete notion of queue length, namely, $Q_i(t) = N_i(t) - S_i(t)$, where $N_i(t)$ is a Poisson process with rate λ, denoting the number of packets arriving at node i during $[0,t]$, while $S_i(t)$ indicates the total number of times node i activates (or deactivates) during $[0,t]$ (we may use λ_U and λ_V to represent different arrival rates for the two sets U and V). The processes $t \mapsto S_i(t)$ and $t \mapsto N_i(t)$ are assumed to be independent. We can define a model where each time a node activates it serves exactly one packet and then deactivates again. The activation clocks still have rates $g_i(Q_i(t))$ with $g_i \in \mathcal{G}$. We can establish results similar to our internal model by adapting the arguments to the discrete setting.

Outline of the chapter. The remainder of this chapter is organized as follows. In Section 2.2 we state large deviation bounds for the input and the output process, which allow us to show that the queue length process at every node has specific lower and upper bounds that hold with very high probability. The proofs of these bounds are deferred to Appendices A–B. In Section 2.3 we use the bounds to couple the lower and the upper external model (with activation rates (2.7) and (2.8), respectively) to the internal model (with activation rates (2.5)). In Section 2.4 we derive the scaling results for the external model, and combine these with the coupling to derive Theorem 2.1.6 (as stated in Section 2.1.2).

§2.2 Bounds for the input and output processes

In this section we state the main results of our analysis of the input process and the output process at a fixed node (recall Definition 1.1.3). With the help of path large deviation techniques, we show that, with high probability as $r \to \infty$, the input process lies in a narrow tube around the deterministic path $t \mapsto (\lambda/\mu)t$ (Proposition 2.2.1). For simplicity, we suppress the index for the arrival rates λ_U and λ_V, and consider a general rate λ. The same holds for $\rho = \lambda/\mu$. We study the output process only for nodes in U, and we give lower and upper bounds in (2.27) and Proposition 2.2.4. We look at a single node and suppress its index, since the queues are independent of each other as long as the nodes remain active or inactive. The proofs of the propositions below for the input process and the output process are given in Appendices A–B, respectively.

Proposition 2.2.1 (Tube for the input process).
For $\delta > 0$ small enough and time horizon $S > 0$, let

$$\Gamma_{S,\delta S} = \left\{ \gamma \in L_\infty([0,S]) \colon \frac{\lambda}{\mu} s - \delta S < \gamma(s) < \frac{\lambda}{\mu} s + \delta S \;\; \forall\, s \in [0,S] \right\}. \tag{2.24}$$

With high probability as $r \to \infty$, the input process lies inside $\Gamma_{S,\delta S}$ as $S \to \infty$, namely,

$$\mathbb{P}\big(Q^+([0,S]) \notin \Gamma_{S,\delta S}\big) = e^{-K_\delta S\,[1+o(1)]}, \qquad S \to \infty. \tag{2.25}$$

$$K_\delta = (\lambda + \delta\mu) + \lambda - 2\sqrt{\lambda(\lambda+\delta\mu)} \in (0,\infty). \tag{2.26}$$

(Note that $\Gamma_{S,\delta S}$ contains negative values. This is of no concern because the path is always non-negative.)

We want to derive lower and upper bounds for the output process for a node in U. An upper bound is trivial by definition, namely,

$$Q^-(t) \le ct, \qquad t \ge 0. \tag{2.27}$$

It is more delicate to compute a lower bound, for which we need some preparatory definitions. We first introduce an auxiliary time that will be useful in our analysis.

Definition 2.2.2 (Auxiliary time).
Consider the internal model and recall that the initial queue lengths at nodes in U are $\gamma_U r$. Define T_U to be the expected time at which the queue length at a node in U hits zero if the transition has not occurred yet. We can write

$$T_U = T_U(r) \sim \alpha r, \qquad r \to \infty, \tag{2.28}$$

with

$$\alpha = \frac{\gamma_U}{c - \rho_U}. \tag{2.29}$$

Note that the quantity αr is the expected time at which the queue length at a node in U hits zero when the node is always active. Since the total inactivity time of a node in U before time T_U will turn out to be negligible compared to αr, we have $T_U \sim \alpha r$ as $r \to \infty$.

Next, we introduce the isolated model, an auxiliary model that will help us to derive a lower bound for the output process. We will see later that the internal model behaves in exactly the same way as the isolated model up to the pre-transition time, in particular, the pre-transition times in the internal and the isolated model coincide in distribution.

Definition 2.2.3 (Isolated model).
In the *isolated model* the activation of nodes in U is not affected by the activity states of nodes in V, i.e., they behave as if they were in isolation. On the other hand, nodes in V are still affected by nodes in U, i.e., they cannot activate until every node in U deactivates. Nodes in V have zero output process.

We study the output process for the isolated model up to time T_U. We will see later in Corollary 2.4.3 that, with high probability as $r \to \infty$, the transition in the internal model occurs before T_U, so it is enough to look at the time interval $[0, T_U]$. In the rare case when the transition does not occur before T_U, we expect it to occur in a very short time after T_U. We are now ready to give the lower bound for the output process.

Proposition 2.2.4 (The output process in the isolated model).
Consider a node in U. For $\delta, \epsilon, \epsilon_1, \epsilon_2 > 0$ small enough, the output process satisfies

$$\begin{aligned}\mathbb{P}_u\big(Q^-(t) < ct - \epsilon r \;\; \forall t \in [0, T_U]\big) \le\, & e^{-K_\delta \alpha r\,[1+o(1)]} + e^{-K_1 r\,[1+o(1)]} \\ & + e^{-\left(K_2 r + K_3 \frac{r}{g_U(r)} + K_4 r \log g_U(r)\right)\,[1+o(1)]}, \qquad r \to \infty,\end{aligned} \tag{2.30}$$

with

$$\begin{aligned} K_1 &= \left(\gamma_U - \frac{2\delta\alpha}{c-\rho_U}\right)\frac{\epsilon_1 - \log(1+\epsilon_1)}{1+\epsilon_1}, \\ K_2 &= \left(\gamma_U - \frac{2\delta\alpha}{c-\rho_U}\right)(1+\epsilon_1)\left(-1-\log\left(\frac{\epsilon_2}{\left(\gamma_U - \frac{2\delta\alpha}{c-\rho_U}\right)(1+\epsilon_1)}\right)\right), \\ K_3 &= \epsilon_2, \\ K_4 &= \left(\gamma_U - \frac{2\delta\alpha}{c-\rho_U}\right)(1+\epsilon_1), \end{aligned} \tag{2.31}$$

satisfying $K_1, K_2, K_3, K_4 \in (0, \infty)$.

By combining the bounds for the input process and the output process, and picking $\delta = \epsilon$ and $S = r$, we obtain lower and upper bounds for the queue length process $Q(t)$ of a node in U.

Corollary 2.2.5 (The queue length process in the isolated model).
For $\delta > 0$ small enough, with high probability as $r \to \infty$, the following bounds hold for a node in U:

$$\begin{aligned} (LB)_U: \quad & Q(t) \ge Q_U^{\mathrm{LB}}(t) = \gamma_U r - (c-\rho_U)t - \delta r, && t \ge 0, \\ (UB)_U: \quad & Q(t) \le Q_U^{\mathrm{UB}}(t) = \gamma_U r - (c-\rho_U)t + 2\delta r, && t \in [0, T_U]. \end{aligned} \tag{2.32}$$

Similarly, with high probability as $r \to \infty$, the following bounds hold for a node in V:

$$\begin{aligned} (LB)_V: \quad & Q(t) \ge Q_V^{\mathrm{LB}}(t) = \gamma_V r + \rho_V t - \delta r, && t \ge 0, \\ (UB)_V: \quad & Q(t) \le Q_V^{\mathrm{UB}}(t) = \gamma_V r + \rho_V t + \delta r, && t \ge 0. \end{aligned} \tag{2.33}$$

Proof. The claim follows from Propositions 2.2.1 and 2.2.4 in combination with the bound in (2.27). □

§2.3 Coupling the internal and the external model

In Sections 2.3.1–2.3.2 we use the bounds defined in Section 2.2 to construct two couplings that allow us to compare the internal and the external model (Proposition 2.3.5, respectively, Proposition 2.3.8 and Corollary 2.3.9). Throughout the sequel we assume that the deactivation rates are fixed, i.e., the deactivation Poisson clocks ring at rate 1. A node can activate only if all its neighbors are inactive. If a node is inactive, then the activation Poisson clocks ring at rates that vary over time in a deterministic way, or as functions of the queue lengths.

We are interested in coupling the models in the time interval $[0, T_U]$ and on the following event.

Definition 2.3.1 (Good behavior).
Let $\mathcal{E}_\delta$ be the event that the queue length processes in the internal model all have *good behavior* in the interval $[0, T_U]$, in the sense that

$$\begin{aligned} \mathcal{E}_\delta = &\big\{Q_U^{\mathrm{LB}}(t) \le Q_i(t) \le Q_U^{\mathrm{UB}}(t)\ \forall t \in [0, T_U]\ \forall i \in U\big\} \\ &\cup \big\{Q_V^{\mathrm{LB}}(t) \le Q_i(t) \le Q_V^{\mathrm{UB}}(t)\ \forall t \in [0, T_U]\ \forall i \in V\big\}, \end{aligned} \tag{2.34}$$

i.e., the paths lie between their respectively lower and upper bounds for nodes in U and V. This event depends on the perturbation parameter δ.

Lemma 2.3.2 (Probability of good behavior).
For $\delta > 0$ small enough,

$$\lim_{r\to\infty} \mathbb{P}_u(\mathcal{E}_\delta) = 1. \tag{2.35}$$

Proof. The claim follows from Corollary 2.2.5. □

In what follows we couple on the event $\mathcal{E}_\delta$ only. The coupling can be extended in an arbitrary way off the event $\mathcal{E}_\delta$. The way this is done is irrelevant because of Lemma 2.3.2.

§2.3.1 Coupling the internal and the lower external model

The lower external model defined in (2.7) can also be described in the following way. At time $t \ge 0$ the activation rates are

$$r_i^{\mathrm{low}}(t) = \begin{cases} g_U(Q_U^{\mathrm{LB}}(t)), & i \in U, \\ g_V(Q_V^{\mathrm{UB}}(t)), & i \in V. \end{cases} \tag{2.36}$$

Note that when the lower bound $Q_U^{\mathrm{LB}}(t)$ becomes negative the activation function g_U is zero by definition. In this way we are able extend the coupling to any time $t \ge 0$, even though we consider only the interval $[0, T_U]$.

Lemma 2.3.3 (Upper bound in the lower external model).
With high probability as $r \to \infty$, the transition time $\mathcal{T}_G^{\mathrm{low}}$ in the lower external model is smaller than T_U, i.e.,

$$\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\mathrm{low}} \le T_U) = 1. \tag{2.37}$$

Proof. As we will see in Section 2.4.2, with high probability as $r \to \infty$, the transition time in the external model is smaller than T_U. Since the lower external model is defined for an arbitrarily small perturbation $\delta > 0$, we conclude by using the continuity of g_U, g_V. □

We introduce a system that allows us to couple the internal model with the lower external model.

Definition 2.3.4 (Coupling system for the lower external model). Suppose that $h_i(t) \geq \max\{Q_U^{\mathrm{UB}}(t), Q_V^{\mathrm{UB}}(t)\}$ for all $i \in U \sqcup V$ and all $t \in [0, T_U]$. Consider a system $\mathcal{H}^{\mathrm{low}}$ where clocks are associated with each node in the following way.

- A Poisson deactivation clock ticks at rate 1. Both the nodes in the lower external model and in the internal model are governed by this clock:
  - if both nodes are active, then they deactivate together;
  - if only one node is active, then it deactivates;
  - if both nodes are inactive, then nothing happens.
- A Poisson activation clock ticks at rate $g_U(h_i(t))$ at time t for a node $i \in U$. Both the nodes in the lower external model and in the internal model are governed by this clock:
  - if both nodes are active, or both are inactive but have active neighbors, then nothing happens;
  - if the node in the internal model is active and the node in the lower external model is not, then the latter node activates (if it can) with probability
$$\frac{r_i^{\mathrm{low}}(t)}{g_U(h_i(t))}; \tag{2.38}$$
  - if both nodes are inactive but can be activated, then this happens with probabilities
$$\begin{aligned} &\frac{r_i^{\mathrm{low}}(t)}{g_U(h_i(t))} \quad \text{for the lower external model,} \\ &\frac{r_i^{\mathrm{int}}(t)}{g_U(h_i(t))} \quad \text{for the internal model,} \end{aligned} \tag{2.39}$$
    where
$$\frac{r_i^{\mathrm{low}}(t)}{g_U(h_i(t))} \leq \frac{r_i^{\mathrm{int}}(t)}{g_U(h_i(t))}, \tag{2.40}$$
    in such a way that if the node in the lower external model activates, then it also activates in the internal model.

- A Poisson activation clock ticks at rate $g_V(h_i(t))$ at time t for a node $i \in V$. The same happens as for the nodes in U, but the activation probabilities are

$$\begin{aligned} &\frac{r_i^{\mathrm{low}}(t)}{g_V(h_i(t))} \quad \text{for the lower external model,} \\ &\frac{r_i^{\mathrm{int}}(t)}{g_V(h_i(t))} \quad \text{for the internal model,} \end{aligned} \tag{2.41}$$

where

$$\frac{r_i^{\mathrm{low}}(t)}{g_U(h_i(t))} \geq \frac{r_i^{\mathrm{int}}(t)}{g_U(h_i(t))}, \tag{2.42}$$

in such a way that if the node in the internal model activates, then it also activates in the lower external model.

With the constructions above, we are now able to compare the transition times of the two models.

Proposition 2.3.5 (Coupling the internal and the lower external model). *The following statements hold.*

(i) Under the coupling $\mathcal{H}^{\mathrm{low}}$, the joint activity processes in the internal and in the lower external model are ordered for all $t \in [0, T_U]$, i.e.,

$$\begin{aligned} X_i^{\mathrm{low}}(t) \leq X_i^{\mathrm{int}}(t), \quad i \in U, \\ X_i^{\mathrm{int}}(t) \leq X_i^{\mathrm{low}}(t), \quad i \in V. \end{aligned} \tag{2.43}$$

(ii) With high probability as $r \to \infty$, the transition time $\mathcal{T}_G^{\mathrm{int}}$ in the internal model is at least as large as the transition time $\mathcal{T}_G^{\mathrm{low}}$ in the lower external model, i.e.,

$$\lim_{r \to \infty} \hat{\mathbb{P}}_u(\mathcal{T}_G^{\mathrm{low}} \leq \mathcal{T}_G^{\mathrm{int}}) = 1, \tag{2.44}$$

where $\hat{\mathbb{P}}_u$ is the joint law induced by the coupling with starting u.

Proof. We prove the two statements separately.

(i) For each node $i \in U$ and for all $t \in [0, T_U]$, we have that $Q_i^{\mathrm{LB}}(t) \leq Q_i(t)$ and $g_U(Q_i^{\mathrm{LB}}(t)) \leq g_U(Q_i(t))$ by the monotonicity of the function g_U. On the other hand, for each node $i \in V$, $Q_i(t) \leq Q_i^{\mathrm{UB}}(t)$ and $g_V(Q_i(t)) \leq g_V(Q_i^{\mathrm{UB}}(t))$ by the monotonicity of the function g_V. Under the system $\mathcal{H}^{\mathrm{low}}$, at any moment the random variable describing the state of a node $i \in U$ in the lower external model is dominated by the one in the internal model, i.e., by (2.40) for all $t \in [0, T_U]$,

$$X_i^{\mathrm{low}}(t) \leq X_i^{\mathrm{int}}(t). \tag{2.45}$$

On the other hand, the random variable describing the state of a node $j \in V$ in the lower external model dominates the one in the internal model, i.e., by (2.42) for all $t \in [0, T_U]$,

$$X_i^{\mathrm{int}}(t) \leq X_i^{\mathrm{low}}(t). \tag{2.46}$$

Hence the joint activity processes in the two models are ordered.

(ii) Using the coupling construction and the ordering above, we can show that, on the event $\mathcal{E}_\delta$, the nodes in U in the lower external model deactivate earlier than in the internal model, and the nodes in V activate earlier in the lower external model. Hence the transition occurs earlier in the lower external model.

Note that we are able to compare the transition times only when $\mathcal{T}_G^{\mathrm{low}} \leq T_U$, so we look at the coupling also on the event $\{\mathcal{T}_G^{\mathrm{low}} \leq T_U\}$, which has high probability as $r \to \infty$ (Lemma 2.3.3). On this event we have $\mathcal{T}_G^{\mathrm{low}} \leq \mathcal{T}_G^{\mathrm{int}}$. Therefore

$$1 = \lim_{r\to\infty} \hat{\mathbb{P}}_u(\mathcal{E}_\delta, \mathcal{T}_G^{\mathrm{low}} \leq T_U, \mathcal{T}_G^{\mathrm{low}} \leq \mathcal{T}_G^{\mathrm{int}}) = \lim_{r\to\infty} \hat{\mathbb{P}}_u(\mathcal{T}_G^{\mathrm{low}} \leq \mathcal{T}_G^{\mathrm{int}}). \tag{2.47}$$

□

§2.3.2 Coupling the isolated and the upper external model

The upper external model defined in (2.8) can also be described in the following way. At time $t \in [0, T_U]$ the activation rates are

$$r_i^{\mathrm{upp}}(t) = \begin{cases} g_U(Q_U^{\mathrm{UB}}(t)), & i \in U, \\ g_V(Q_V^{\mathrm{LB}}(t)), & i \in V. \end{cases} \tag{2.48}$$

Lemma 2.3.6 (Upper bound in the upper external model).
With high probability as $r \to \infty$, the transition time $\mathcal{T}_G^{\mathrm{upp}}$ in the upper external model is smaller than T_U, i.e.,

$$\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\mathrm{upp}} \leq T_U) = 1. \tag{2.49}$$

This statement is to be read as follows. Let δ be the perturbation parameter in the upper external model appearing in (2.8)*. Then for every $\delta > 0$ there exists a $\delta'(\delta) > 0$, satisfying* $\lim_{\delta\to 0} \delta'(\delta) = 0$*, such that* $\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\mathrm{upp}} \leq [1 + \delta'(\delta)]T_U) = 1$.

Proof. Analogous to the proof of Lemma 2.3.3. □

We introduce a system that allows us to couple the isolated model with the upper external model up to time τ_G^{iso}.

Definition 2.3.7 (Coupling system for the upper external model).
Suppose that $h_i(t) \geq \max\{Q_U^{\mathrm{UB}}(t), Q_V^{\mathrm{UB}}(t)\}$ for all $i \in U \sqcup V$ and all $t \in [0, \tau_G^{\mathrm{iso}}]$. Couple the processes in the same way as in Definition 2.3.4 for $\mathcal{H}^{\mathrm{low}}$, but with different activation probabilities. The probabilities for the isolated model and for the upper external model are such that

$$\begin{aligned} \frac{r_i^{\mathrm{iso}}(t)}{g_U(h_i(t))} &\leq \frac{r_i^{\mathrm{upp}}(t)}{g_U(h_i(t))}, \quad i \in U, \\ \frac{r_i^{\mathrm{upp}}(t)}{g_V(h_i(t))} &\leq \frac{r_i^{\mathrm{iso}}(t)}{g_V(h_i(t))}, \quad i \in V, \end{aligned} \tag{2.50}$$

where for $t \in [0, \tau_G^{\text{iso}}]$

$$r_i^{\text{iso}}(t) = \begin{cases} g_U(Q_i(t)), & i \in U, \\ g_V(Q_i(t)), & i \in V. \end{cases} \tag{2.51}$$

Note that when $\tau_G^{\text{iso}} \leq T_U$, the isolated model behaves exactly as the internal model in the interval $[0, \tau_G^{\text{iso}}]$, as shown in Appendix B.2. Moreover, the coupling is defined only when $\tau_G^{\text{iso}} \leq T_U$. We look then at the coupling also on the event $\{\mathcal{T}_G^{\text{upp}} \leq T_U\}$, which has high probability as $r \to \infty$ (Lemma 2.3.6). In the following proposition we see how this ensures that the coupling is well defined, and we compare the pre-transition times of the two models.

Proposition 2.3.8 (Coupling the isolated and the upper external model).
The following statements hold.

(i) Under the coupling $\mathcal{H}^{\text{upp}}$, the joint activity processes in the isolated model and in the upper external model are ordered up to time τ_G^{iso}, i.e., for all $t \in [0, \tau_G^{\text{iso}}]$,

$$\begin{aligned} X_i^{\text{iso}}(t) &\leq X_i^{\text{upp}}(t), \quad i \in U, \\ X_i^{\text{upp}}(t) &\leq X_i^{\text{iso}}(t), \quad i \in V. \end{aligned} \tag{2.52}$$

(ii) With high probability as $r \to \infty$, the pre-transition time τ_G^{upp} in the upper external model is at least as large as the pre-transition time τ_G^{iso} in the isolated model, i.e.,

$$\lim_{r \to \infty} \hat{\mathbb{P}}_u(\tau_G^{\text{iso}} \leq \tau_G^{\text{upp}}) = 1, \tag{2.53}$$

where $\hat{\mathbb{P}}_u$ is the joint law induced by the coupling with starting u.

Proof. We prove the two statements separately.

(i) The proof is analogous to that of Proposition 2.3.5, but this time we use the system $\mathcal{H}^{\text{upp}}$ up to time τ_G^{iso} and all the inequalities are reversed.

(ii) Using the coupling construction and the ordering above, we can show that, on the event $\mathcal{E}_\delta \cap \{\mathcal{T}_G^{\text{upp}} \leq T_U\}$, the nodes in U in the isolated model deactivate earlier than in the upper external model, and the first activating node in V activates earlier in the isolated model. Hence the pre-transition occurs earlier in the isolated model, and we have $\tau_G^{\text{iso}} \leq \tau_G^{\text{upp}} \leq \mathcal{T}_G^{\text{upp}} \leq T_U$. Therefore the coupling is well defined and

$$1 = \lim_{r \to \infty} \hat{\mathbb{P}}_u(\mathcal{E}_{\delta, T_U}, \mathcal{T}_G^{\text{upp}} \leq T_U, \tau_G^{\text{iso}} \leq \tau_G^{\text{upp}}) = \lim_{r \to \infty} \hat{\mathbb{P}}_u(\tau_G^{\text{iso}} \leq \tau_G^{\text{upp}}). \tag{2.54}$$

□

Corollary 2.3.9 (Comparing times between models).
With high probability as $r \to \infty$, the transition time $\mathcal{T}_G^{\text{upp}}$ in the upper external model is at least as large as the pre-transition time τ_G^{int} in the internal model, i.e.,

$$\lim_{r \to \infty} \hat{\mathbb{P}}_u(\tau_G^{\text{int}} \leq \mathcal{T}_G^{\text{upp}}) = 1. \tag{2.55}$$

Proof. Since $\lim_{r\to\infty}\mathbb{P}(\tau_G^{\mathrm{iso}} \le T_U) = 1$, we have, as shown in Proposition B.6 in Appendix B.2, that the pre-transition times in the isolated model and in the internal model coincide. Hence

$$1 = \lim_{r\to\infty}\hat{\mathbb{P}}_u(\tau_G^{\mathrm{iso}} \le \tau_G^{\mathrm{upp}}) = \lim_{r\to\infty}\hat{\mathbb{P}}_u(\tau_G^{\mathrm{int}} \le \tau_G^{\mathrm{upp}}) \le \lim_{r\to\infty}\hat{\mathbb{P}}_u(\tau_G^{\mathrm{int}} \le \mathcal{T}_G^{\mathrm{upp}}), \tag{2.56}$$

which completes the proof. □

§2.4 Proofs of the main results

The goal of this section is to identify the asymptotic behavior of the transition time in the internal model. In Sections 2.4.1–2.4.2 we look at the external model and prove Theorems 2.1.4–2.1.5, respectively. In Section 2.4.3 we show that the difference between the transition time and the pre-transition time is negligible for all the models considered. In Section 2.4.4 we put these results together to prove Theorem 2.1.6.

§2.4.1 Proof: critical time scale in the external model

In this section we prove Theorem 2.1.4. From now on we write $a(r) \sim b(r)$ to indicate that $\lim_{r\to\infty} a(r)/b(r) = 1$, while we write $a(r) \asymp b(r)$ to indicate that $0 < \liminf_{r\to\infty} a(r)/b(r) \le \limsup_{r\to\infty} a(r)/b(r) < \infty$.

Proof of Theorem 2.1.4. In order to compute the critical time scale M_c, we must solve the equation $M\nu(M) = 1$ in (2.12). We know from (2.9) and (2.13) that

$$\nu(s) \sim |U| r_U(s)^{1-|U|}, \qquad r \to \infty. \tag{2.57}$$

We want to identify how the transition time is related to the choice of g_U in Definition 2.1.1. Consider the time scale $M_c = F_c r^\gamma$, where $\gamma \in (0,1]$ and $F_c \in (0,\infty)$. As $r \to \infty$, we have

$$\begin{aligned} 1 = r^0 &= M_c\nu(M_c) = F_c r^\gamma\, \nu(F_c r^\gamma) \sim F_c r^\gamma |U| r_U(F_c r^\gamma)^{-(|U|-1)} \\ &= F_c r^\gamma |U| g_U\big(\gamma_U r - (c-\rho_U)F_c r^\gamma\big)^{-(|U|-1)} \\ &\sim F_c r^\gamma |U| B^{-(|U|-1)}\big(\gamma_U r - (c-\rho_U)F_c r^\gamma\big)^{-\beta(|U|-1)}. \end{aligned} \tag{2.58}$$

Recall from (2.29) that $\alpha = \frac{\gamma_U}{c-\rho_U}$. We distinguish between three cases.

(I) Case $\gamma \in (0,1)$ and $F_c \in (0,\infty)$. As $r \to \infty$, the criterion in (2.58) reads

$$1 = r^0 \sim F_c r^\gamma |U| B^{-(|U|-1)} (\gamma_U r)^{-\beta(|U|-1)}. \tag{2.59}$$

In order for the exponents of r to match, we need

$$\beta = \frac{\gamma}{|U|-1}. \tag{2.60}$$

Inserting (2.60) into (2.59), we get

$$F_c |U| B^{-(|U|-1)} \gamma_U^{-\beta(|U|-1)} = 1, \tag{2.61}$$

which gives

$$F_c = \frac{\gamma_U^{\beta(|U|-1)} B^{(|U|-1)}}{|U|}. \tag{2.62}$$

Hence

$$M_c = \frac{(B\gamma_U^{\beta})^{|U|-1}}{|U|} r^{\beta(|U|-1)}, \qquad r \to \infty. \tag{2.63}$$

(II) Case $\gamma = 1$ and $F_c \in (0, \alpha)$. As $r \to \infty$, the criterion in (2.58) reads

$$1 = r^0 \sim F_c |U| B^{-(|U|-1)} (\gamma_U - (c - \rho_U) F_c)^{-\beta(|U|-1)} r^{1-\beta(|U|-1)}. \tag{2.64}$$

In order for the exponents of r to match, we need

$$\beta = \frac{1}{|U|-1}. \tag{2.65}$$

Inserting (2.65) into (2.64), we get

$$\frac{F_c |U| B^{-(|U|-1)}}{\gamma_U - (c - \rho_U) F_c} = 1, \tag{2.66}$$

which gives

$$F_c = \frac{\gamma_U}{|U| B^{-(|U|-1)} + (c - \rho_U)}. \tag{2.67}$$

Hence

$$M_c = \frac{\gamma_U}{|U| B^{-(|U|-1)} + (c - \rho_U)} r, \qquad r \to \infty. \tag{2.68}$$

Recall from (2.28) that $T_U \sim \alpha r$ is the expected time at which the queue length at a node in U hits zero. We will see in Section 2.4.2 that the transition in the external model typically occurs before the queues are empty.

(III) Case $\gamma = 1$ and $F_c = \alpha - D r^{-\delta}$, $\delta \in (0, 1)$. As $r \to \infty$, the criterion in (2.58) reads

$$1 = r^0 \sim \alpha r |U| B^{-(|U|-1)} \big((c - \rho_U) D r^{1-\delta}\big)^{-\beta(|U|-1)}. \tag{2.69}$$

In order for the exponents of r to match, we need

$$\beta = \frac{1}{(1-\delta)(|U|-1)}. \tag{2.70}$$

Inserting (2.70) into (2.69), we get

$$\alpha |U| B^{-(|U|-1)} ((c - \rho_U) D)^{-\beta(|U|-1)} = 1, \tag{2.71}$$

which gives

$$D = \frac{(\alpha |U| B^{-(|U|-1)})^{1/\beta(|U|-1)}}{c - \rho_U}. \tag{2.72}$$

Hence

$$M_c = \alpha r - \frac{(\alpha |U| B^{-(|U|-1)})^{1/\beta(|U|-1)}}{c - \rho_U} r^{1/\beta(|U|-1)} = \alpha r, \qquad r \to \infty, \tag{2.73}$$

and so the crossover takes place in a window of size $O(r^{1/\beta(|U|-1)}) = o(r)$ around αr. Note that this window gets narrower as β increases, i.e., as the activation rates of nodes in U increase.

□

In the following remark we discuss more general activation functions g_U.

Remark 2.4.1 (Modulation with slowly varying functions).
Consider activation functions of the form $g_U(x) = x^\beta \hat{\mathcal{L}}(x)$ with $\beta \in (0,\infty)$ and $\hat{\mathcal{L}}(x)$ a slowly varying function (i.e., $\lim_{x\to\infty} \hat{\mathcal{L}}(ax)/\hat{\mathcal{L}}(x) = 1$ for all $a > 0$). Let $M_c = r^\gamma \mathcal{L}(r)$ with $\gamma \in (0,1)$ and $\mathcal{L}(r)$ a slowly varying function. As $r \to \infty$, we have

$$\begin{aligned} 1 = r^0 &\sim r^\gamma \mathcal{L}(r)|U|\big(\gamma_U r - (c-\rho_U) r^\gamma \mathcal{L}(r)\big)^{-\beta(|U|-1)} \hat{\mathcal{L}}(\gamma_U r - (c-\rho_U) r^\gamma \mathcal{L}(r))^{-(|U|-1)} \\ &\sim r^\gamma \mathcal{L}(r)|U|(\gamma_U r)^{-\beta(|U|-1)} \hat{\mathcal{L}}(\gamma_U r)^{-(|U|-1)}. \end{aligned} \tag{2.74}$$

In order for the exponents of r to match, we again need

$$\beta = \frac{\gamma}{|U|-1}. \tag{2.75}$$

We get

$$\mathcal{L}(r) = \frac{\gamma_U^{\beta(|U|-1)}}{|U|} \hat{\mathcal{L}}(\gamma_U r)^{|U|-1}, \qquad r \to \infty. \tag{2.76}$$

Hence

$$M_c = \frac{\gamma_U^{\beta(|U|-1)}}{|U|} r^{\beta(|U|-1)} \hat{\mathcal{L}}(\gamma_U r)^{|U|-1}, \qquad r \to \infty. \tag{2.77}$$

We can even include the case $\beta = 0$, in which we obtain that if $g_U(x) = \hat{\mathcal{L}}(x)$ with $\lim_{x\to\infty} \hat{\mathcal{L}}(x) = \infty$, then

$$M_c = \frac{1}{|U|} \hat{\mathcal{L}}(\gamma_U r)^{|U|-1}, \qquad r \to \infty. \tag{2.78}$$

§2.4.2 Proof: transition time in the external model

In this section we prove Theorem 2.1.5. We already know that the transition occurs on the critical time scale M_c computed in Section 2.4.2.

Proof of Theorem 2.1.5. Knowing the critical time scale M_c, we can compute the mean transition time from (2.11). As $r \to \infty$, we have

$$\begin{aligned} \mathbb{E}_u[\mathcal{T}_G^{\text{ext}}] &= \int_0^\infty \mathbb{P}_u(\mathcal{T}_G^{\text{ext}} > x)\, dx = M_c \int_0^\infty \mathbb{P}_u\left(\frac{\mathcal{T}_G^{\text{ext}}}{M_c} > x\right) dx \\ &\sim M_c \int_0^\infty e^{-\int_0^x M_c \nu(M_c s)\, ds}\, dx = M_c \int_0^\infty e^{-\int_0^x \frac{M_c \nu(M_c s)}{M_c \nu(M_c)}\, ds}\, dx \\ &= M_c \int_0^\infty e^{-\int_0^x \left(\frac{\gamma_U r - (c-\rho_U) M_c s}{\gamma_U r - (c-\rho_U) M_c}\right)^{-\beta(|U|-1)} ds}\, dx, \end{aligned} \tag{2.79}$$

where the choice of β is important. We distinguish between the three cases.

(I) Case $\beta \in (0, \frac{1}{|U|-1})$, $M_c = F_c r^\gamma$, $\gamma \in (0,1)$. We have

$$\lim_{r\to\infty} \left(\frac{\gamma_U r - (c-\rho_U) M_c s}{\gamma_U r - (c-\rho_U) M_c} \right)^{-\beta(|U|-1)} = 1. \tag{2.80}$$

Hence, as $r \to \infty$,

$$\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}] \sim M_c \int_0^\infty e^{-\int_0^x ds}\, dx = M_c \int_0^\infty e^{-x}\, dx = M_c. \tag{2.81}$$

The law of $\mathcal{T}_G^{\text{ext}}$ is exponential, i.e.,

$$\lim_{r\to\infty} \mathbb{P}_u \left(\frac{\mathcal{T}_G^{\text{ext}}}{\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}]} > x \right) = e^{-x}, \qquad x \in [0,\infty). \tag{2.82}$$

(II) Case $\beta = \frac{1}{|U|-1}$, $M_c = F_c r$, $F_c \in (0,\alpha)$. We have

$$\begin{aligned}\lim_{r\to\infty} \left(\frac{\gamma_U r - (c-\rho_U) F_c r s}{\gamma_U r - (c-\rho_U) F_c r} \right)^{-\beta(|U|-1)} &= \frac{\gamma_U - (c-\rho_U) F_c}{\gamma_U - (c-\rho_U) F_c s} = \frac{1 - \frac{c-\rho_U}{\gamma_U} F_c}{1 - \frac{c-\rho_U}{\gamma_U} F_c s} \\ &= \frac{1 - \frac{F_c}{\alpha}}{1 - \frac{F_c}{\alpha} s} = \frac{1-C}{1-Cs},\end{aligned} \tag{2.83}$$

with $C = F_c/\alpha$. Hence, as $r \to \infty$,

$$\begin{aligned}\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}] &\sim M_c \int_0^{\frac{1}{C}} e^{-\int_0^x \frac{1-C}{1-Cs} ds}\, dx = M_c \int_0^{\frac{1}{C}} e^{-\log(1-Cx)^{-\frac{1-C}{C}}}\, dx \\ &= M_c \int_0^{\frac{1}{C}} (1-Cx)^{\frac{1-C}{C}}\, dx = M_c \left[(1-Cx)^{1+\frac{1-C}{C}} \frac{1}{(1+\frac{1-C}{C})(-C)} \right]_0^{\frac{1}{C}} \\ &= M_c \left[-(1-Cx)^{\frac{1}{C}} \right]_0^{\frac{1}{C}} = M_c.\end{aligned} \tag{2.84}$$

Here, the integral must be truncated at $x = 1/C$ because for larger x the integrand becomes negative. Indeed, note that when $x = 1/C = \alpha/F_c$, which corresponds to time $T_U = \alpha r$, we have

$$\begin{aligned}\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\text{ext}} > T_U) &= \lim_{r\to\infty} \mathbb{P}_u\left(\mathcal{T}_G^{\text{ext}} > \frac{\alpha}{F_c} F_c r \right) = \lim_{r\to\infty} \mathbb{P}_u \left(\frac{\mathcal{T}_G^{\text{ext}}}{M_c} > \frac{\alpha}{F_c} \right) \\ &= \left(1 - C\frac{\alpha}{F_c} \right)^{\frac{1-C}{C}} = 0,\end{aligned} \tag{2.85}$$

because $C = F_c/\alpha$. This means that, with high probability as $r \to \infty$, the transition occurs before time T_U. The law of $\mathcal{T}_G^{\text{ext}}$ is truncated polynomial:

$$\lim_{r\to\infty} \mathbb{P}_u \left(\frac{\mathcal{T}_G^{\text{ext}}}{\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}]} > x \right) = \begin{cases} (1-Cx)^{\frac{1-C}{C}}, & x \in [0, \frac{1}{C}), \\ 0, & x \in [\frac{1}{C}, \infty). \end{cases} \tag{2.86}$$

(III) Case $\beta \in (\frac{1}{|U|-1}, \infty)$, $M_c = \alpha r$. This case corresponds to the limit $C \to 1$ of the previous case. In this limit, (2.86) becomes

$$\lim_{r\to\infty} \mathbb{P}_u\left(\frac{\mathcal{T}_G^{\text{ext}}}{\mathbb{E}_u[\mathcal{T}_G^{\text{ext}}]} > x\right) = \begin{cases} 1, & x \in [0,1), \\ 0, & x \in [1,\infty). \end{cases} \tag{2.87}$$

□

Note that the three cases above corresponds to the three regimes of behavior: respectively, the subcritical regime, the critical regime and the supercritical regime.

§2.4.3 Negligible gap in the internal model

In this section we focus on the internal model and estimate the length of the interval $[\tau_G^{\text{int}}, \mathcal{T}_G^{\text{int}}]$, which, with high probability as $r \to \infty$, turns out to be very small with respect to τ_G^{int}. This implies that the transition time has the same asymptotic behavior as the pre-transition time.

We know that the queue at a node $i \in V$ is of order r at time τ_G^{int}, i.e., $Q_i(\tau_G^{\text{int}}) \asymp r$, since it starts at $\gamma_V r$, with $\gamma_V > 0$, and only the input process is present until this time. Hence all the activation Poisson clocks at nodes in V tick at a very aggressive rate. The idea is that within the activation period (which has an exponential distribution with mean 1) of the first node activating in V, all the other nodes in V activate because they are not “blocked" by any node in U. Consequently, the network quickly reaches v.

Theorem 2.4.2 (Negligible gap).
In the internal model

$$\lim_{r\to\infty} \mathbb{P}_u\left(\mathcal{T}_G^{\text{int}} - \tau_G^{\text{int}} = o\left(\frac{1}{g_V(r)}\right)\right) = 1. \tag{2.88}$$

Proof. Starting from τ_G^{int}, a node $x \in V$ remains inactive for an exponential period with mean $1/r_x^{\text{int}}(\tau_G) = 1/g_V(Q(\tau_G)) \asymp 1/g_V(r)$. Denote by W_x the length of an inactivity period for a node $x \in V$. Let x_1 be the first node activating in V. We then have i.i.d. inactivity periods $W_x \simeq \text{Exp}(g_V(Q(\tau_G)))$ for all $x \in V \setminus \{x_1\}$. We label the remaining nodes $x_2, \dots, x_{|V|}$ in an arbitrary way. We also have i.i.d. activity periods $Z_x \simeq \text{Exp}(1)$ for all $x \in V$.

Consider a time $t_1 = o(1/g_V(r))$. With high probability as $r \to \infty$, all the nodes in V activate before time t_1, i.e.,

$$\begin{aligned} \lim_{r\to\infty} \mathbb{P}_u\big(W_{x_i} < t_1, \forall i = 2, \dots, |V|\big) &= \lim_{r\to\infty} \mathbb{P}_u\big(W_{x_2} < t_1\big)^{|V|-1} \\ &= \lim_{r\to\infty} \big(1 - e^{-g_V(Q(\tau_G))t_1}\big)^{|V|-1} = 1. \end{aligned} \tag{2.89}$$

Moreover, with high probability as $r \to \infty$, once they activated, all nodes in V stay active for a period of length at least $t_2 \asymp 1/g_V(r) > t_1$, i.e.,

$$\begin{aligned} \lim_{r\to\infty} \mathbb{P}_u\big(Z_{x_i} > t_2 \;\; \forall i = 1, \dots, |V|\big) &= \lim_{r\to\infty} \mathbb{P}_u\big(Z_{x_1} > t_2\big)^{|V|} \\ &= \lim_{r\to\infty} (e^{-t_2})^{|V|} = 1. \end{aligned} \tag{2.90}$$

In conclusion, with high probability as $r \to \infty$, every node in V activates before time t_1 and remains active for at least a time $t_2 > t_1$. This ensures that the transition occurs before time t_2. In particular, it occurs when the last node in V activates (which occurs even before time t_1), so that $\mathcal{T}_G^{\text{int}} - \tau_G^{\text{int}} = o(1/g_V(r))$. □

Note that this argument extends to any "external" model with activation rates that tend to infinity with r, in particular, to all the models considered in the chapter. The transition always happens quickly after the pre-transition, due to the high level of aggressiveness of nodes in V.

Corollary 2.4.3 (Upper bound on the transition time).
With high probability as $r \to \infty$, the transition time in the internal model is smaller than T_U, i.e.,

$$\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\text{int}} \le T_U) = 1. \tag{2.91}$$

Proof. The claim follows from Lemma 2.3.6, Corollary 2.3.9 and Theorem 2.4.2. □

§2.4.4 Proof: transition time in the internal model

In this section we prove Theorem 2.1.6. First we derive the sandwich of the transition times in the lower external, the internal and the upper external model. After that we identify the asymptotics of the transition time for the internal model by using the results for the external models.

Proof of Theorem 2.1.6. Using Proposition 2.3.5, Corollary 2.3.9 and Theorem 2.4.2, we have that there exists a coupling such that

$$\begin{aligned} 1 &= \lim_{r\to\infty} \hat{\mathbb{P}}_u\left(\mathcal{T}_G^{\text{low}} \le \mathcal{T}_G^{\text{int}}, \mathcal{T}_G^{\text{int}} = \tau_G^{\text{int}} + o\left(\frac{1}{g_V(r)}\right), \tau_G^{\text{int}} \le \mathcal{T}_G^{\text{upp}}\right) \\ &= \lim_{r\to\infty} \hat{\mathbb{P}}_u\left(\mathcal{T}_G^{\text{low}} \le \mathcal{T}_G^{\text{int}} \le \mathcal{T}_G^{\text{upp}} + o\left(\frac{1}{g_V(r)}\right)\right) \\ &= \lim_{r\to\infty} \hat{\mathbb{P}}_u(\mathcal{T}_G^{\text{low}} \le \mathcal{T}_G^{\text{int}} \le \mathcal{T}_G^{\text{upp}}), \end{aligned} \tag{2.92}$$

where $\hat{\mathbb{P}}_u$ is the joint law of the three models on the same probability space all three starting from u.

By Theorem 2.1.5, we know the law of the transition time in the external model. By construction, we have $\mathbb{E}_u[\mathcal{T}_G^{\text{low}}] \le \mathbb{E}_u[\mathcal{T}_G^{\text{ext}}] \le \mathbb{E}_u[\mathcal{T}_G^{\text{upp}}]$. When considering the lower and the upper external model, the transition time asymptotics are controlled by the prefactors $F_{c,\delta}^{\text{low}}$ and $F_{c,\delta}^{\text{upp}}$, respectively, which are perturbations of the prefactor F_c due to the perturbations of the activation rates. In particular, we know from (2.20) that $\lim_{\delta\to 0} F_{c,\delta}^{\text{low}} = \lim_{\delta\to 0} F_{c,\delta}^{\text{upp}} = F_c$. Hence, for all $\epsilon > 0$,

$$\mathbb{E}_u[\mathcal{T}_G^{\text{int}}] = (F_c \pm \epsilon) r^{\beta(|U|-1)} [1 + o(1)], \qquad r \to \infty, \tag{2.93}$$

and since ϵ can be taken arbitrarily small, it may be absorbed into the $o(1)$-term, as

$$\mathbb{E}_u[\mathcal{T}_G^{\text{int}}] = F_c r^{\beta(|U|-1)} [1 + o(1)], \qquad r \to \infty. \tag{2.94}$$

The same kind of argument applies to the law of the transition time, since for any $x \in [0, \infty)$,

$$\lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\text{low}} > x) \le \lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\text{int}} > x) \le \lim_{r\to\infty} \mathbb{P}_u(\mathcal{T}_G^{\text{upp}} > x). \tag{2.95}$$

This completes the proof. □

§A Appendix: the input process

The main target of this appendix is to prove Proposition 2.2.1 in Section 2.2. We use path large deviation techniques. For simplicity, we suppress the index for the arrival rates λ_U and λ_V, and consider a general rate λ. We show that, with high probability as $r \to \infty$, the input process lies in a narrow tube around the deterministic path $t \mapsto (\lambda/\mu)t$.

Consider a single queue, and for simplicity suppress its index. For $T > 0$, define the scaled process

$$Q_n^+(t) = \frac{1}{n} Q^+(nt) = \frac{1}{n} \sum_{j=1}^{N(nt)} Y_j, \qquad t \in [0, T], \tag{2.96}$$

with $Q_n^+(0) = 0$. We have

$$\mathbb{E}[Q_n^+(t)] = \frac{1}{n}\frac{\lambda nt}{\mu} = \frac{\lambda}{\mu} t, \tag{2.97}$$

and, by the strong law of large numbers, $Q_n^+(t) \to (\lambda/\mu)t$ almost surely for every t as $n \to \infty$.

When studying the process $t \mapsto Q_n^+(t)$, we need to take into account that this is a combination of the Poisson arrival process $t \mapsto N(nt)$ and the exponential service times Y_j, $j \in \mathbb{N}$. Two different types of fluctuations can occur: packets arrive at a slower or faster rate than λ, respectively, service times for each packet are shorter or longer than their mean $1/\mu$. Both need to be considered for a proper large deviation analysis.

§A.1 Large deviation principle for the two components

Definition A.1 (Space of paths).
Consider the space $L_\infty([0,T])$ of *essentially bounded* functions in $[0,T]$, with the norm $\|f\|_\infty = \operatorname{ess\,sup}_{x\in[0,T]} |f(x)|$ called the essential norm. A function f is essentially bounded, i.e., $f \in L_\infty([0,T])$, when there is a measurable function g on $[0,T]$ such that $f = g$ except on a set of measure zero and g is bounded. Let $\mathcal{AC}_T \subset L_\infty([0,T])$ denote the space of *absolutely continuous* functions $f\colon [0,T] \to \mathbb{R}$ such that $f(0) = 0$.

Given the Poisson arrival process $t \mapsto N(nt)$ with rate λ, define the scaled process $t \mapsto Z_n(t)$ by

$$Z_n(t) = \frac{1}{n} N(nt) = \frac{1}{n} \sum_{i=1}^{nt} X_i = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \qquad t \in [0,T], \tag{2.98}$$

where $X_i \simeq \text{Poisson}(\lambda)$ are i.i.d. random variables and $\lfloor x \rfloor$ denotes the greatest integer smaller than or equal to x. Note that $N(nt) \simeq \text{Poisson}(\lambda nt)$. Let ν_n be the law of $(Z_n(t))_{t\in[0,T]}$ on $L_\infty([0,T])$. Note that $Z_n(t)$ is asymptotically equivalent to $N(t)$ with mean $\mathbb{E}[Z_n(t)] = \lambda t$, and $(Z_n(t))_{t\in[0,T]}$ tends to $(\lambda t)_{t\in[0,T]}$ as $n \to \infty$.

We recall the definition of large deviation principle.

Definition A.2 (Large deviation principle (LDP)).
A family of probability measures $(P_n)_{n\in\mathbb{N}}$ on a Polish space $\mathcal{X}$ is said to satisfy the *large deviation principle (LDP)* with rate n and with good rate function $I\colon \mathcal{X} \to [0,\infty]$ if

$$\begin{aligned} \limsup_{n\to\infty} \frac{1}{n} \log P_t(C) &\leq -I(C) \qquad \forall\, C \subset \mathcal{X} \text{ closed}, \\ \liminf_{n\to\infty} \frac{1}{n} \log P_t(O) &\geq -I(O) \qquad \forall\, O \subset \mathcal{X} \text{ open}, \end{aligned} \tag{2.99}$$

where $I(S) = \inf_{x\in S} I(x)$, $S \subset \mathcal{X}$. A good rate function satisfies: (1) $I \not\equiv \infty$, (2) I is lower semi-continuous, (3) I has compact level sets.

We begin by stating the LDP for the arrival process $(Z_n(t))_{t\in[0,T]}$.

Lemma A.3 (LDP for the arrival process).
The family of probability measures $(\nu_n)_{n\in\mathbb{N}}$ satisfies the LDP on $L_\infty([0,T])$ with rate n and with good rate function I_N given by

$$I_N(\eta) = \begin{cases} \int_0^T \Lambda_N^*(\dot\eta(t))\, dt, & \eta \in \mathcal{AC}_T, \\ \infty, & \textit{otherwise}, \end{cases} \tag{2.100}$$

where $\Lambda_N^*(x) = x \log(x/\lambda) - x + \lambda$, $x \in (0,\infty)$.

Proof. Apply Mogulskii's theorem (see [40, Theorem 5.1.2]). Use the fact that Λ_N^* is the Fenchel-Legendre transform of the cumulant generating function Λ defined by $\Lambda(\theta) = \log \mathbb{E}(e^{\theta X_1})$, $\theta \in \mathbb{R}$. □

For $\Gamma \subset L_\infty([0,T])$, define $I_N(\Gamma) = \inf_{\eta\in\Gamma} I_N(\eta)$. Consequently, the LDP implies that, if $\Gamma \subset L_\infty([0,T])$ is an I_N-continuous set, i.e., $I_N(\Gamma) = I_N(\text{int}(\Gamma)) = I_N(\text{cl}(\Gamma))$, then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Z_n([0,T]) \in \Gamma\big) = -I_N(\Gamma). \tag{2.101}$$

Informally, the LDP reads as the approximate statement

$$\mathbb{P}\big(Z_n([0,T]) \approx \eta([0,T]\big) = e^{-nI_N(\eta)[1+o(1)]}, \qquad n \to \infty, \tag{2.102}$$

where $\approx$ stands for close in the essential norm. Informally, on this event we may approximate

$$Q_n^+(t) = \frac{1}{n}\sum_{j=1}^{N(nt)} Y_j = \frac{1}{n}\sum_{j=1}^{nZ_n(t)} Y_j \approx \frac{1}{n}\sum_{j=1}^{n\eta(t)} Y_j = \frac{1}{n}\sum_{j=1}^{\lfloor n\eta(t)\rfloor} Y_j, \qquad t \in [0,T], \tag{2.103}$$

where $\approx$ now stands for close in the Euclidean norm. Given $\eta \in L_\infty([0,T])$, let μ_n^η denote the law of the last sum in (2.103). Below we state the LDP for the input process subject to the arrival process.

Lemma A.4 (LDP for the input process subject to the arrival process).
Given $\eta \in L_\infty([0,T])$, the family of probability measures $(\mu_n^\eta)_{n\in\mathbb{N}}$ satisfies the LDP on $L_\infty([0,T])$ with rate n and with good rate function I_Q^η given by

$$I_Q^\eta(\phi) = \begin{cases} \int_0^T \Lambda_Q^*\left(\frac{d\phi(t)}{d\eta(t)}\right) d\eta(t), & \phi \in \mathcal{AC}_T, \\ \infty, & \textit{otherwise}, \end{cases} \tag{2.104}$$

where $\Lambda_Q^*(x) = x\mu - 1 - \log(x\mu)$, $x \in (0,\infty)$.

Proof. Again apply Mogulskii's theorem, this time with $\eta(t)$ as the time index. Use that Λ^* is the Fenchel-Legendre transform of the cumulant generating function Λ defined by $\Lambda(\theta) = \log \mathbb{E}(e^{\theta Y_1})$, $\theta \in \mathbb{R}$. □

§A.2 Measures in product spaces

The rate function I_Q^η describes the large deviations for the sequence of processes $(Q_n^+(t))_{t\in[0,T]}$ given the path η. To derive the LDP averaged over η, we need a small digression into measures in product spaces.

Definition A.5 (Product measures).
Define the family of probability measures $(\rho_n)_{n\in\mathbb{N}}$ such that $\rho_n = \nu_n\mu_n^\eta$. These measures are defined on the product space $L_\infty([0,T]) \times L_\infty([0,T])$ given by the Cartesian product of the space $L_\infty([0,T])$ with itself, equipped with the product topology.

The open sets in the product topology are unions of sets of the form $U_1 \times U_2$ with U_1, U_2 open in $L_\infty([0,T])$. Moreover, the product of base elements of $L_\infty([0,T])$ gives a basis for the product space $L_\infty([0,T]) \times L_\infty([0,T])$. Define the projections $\mathrm{Pr}_i\colon L_\infty([0,T]) \times L_\infty([0,T]) \to L_\infty([0,T])$, $i = 1,2$, onto the first and the second coordinates, respectively. The product topology on $L_\infty([0,T]) \times L_\infty([0,T])$ is the topology generated by sets of the form $\mathrm{Pr}_i^{-1}(U_i)$, $i = 1,2$, where and U_1, U_2 are open subsets of $L_\infty([0,T])$.

Lemma A.6 (Product LDP).
The family of probability measures $(\rho_n)_{n\in\mathbb{N}}$ satisfies the LDP on $L_\infty([0,T])\times L_\infty([0,T])$ with rate n and with good rate function I given by

$$I(\phi,\eta) = \begin{cases} \int_0^T \Lambda_Q^*\left(\frac{d\phi(t)}{d\eta(t)}\right) d\eta(t) + \int_0^T \Lambda_N^*(\dot\eta(t))\, dt, & \phi,\eta \in \mathcal{AC}_T, \\ \infty, & \textit{otherwise}. \end{cases} \tag{2.105}$$

Proof. The claim follows from standard large deviation theory (see [40]). □

§A.3 Large deviation principle for the input process

The contraction principle allows us to derive the LDP averaged over η. Indeed, let $\mathcal{X} = L_\infty([0,T]) \times L_\infty([0,T])$ and $\mathcal{Y} = L_\infty([0,T])$, let $(\rho_n)_{n\in\mathbb{N}}$ be a sequence of product measures on $\mathcal{X}$, and consider the projection Pr_1 onto $\mathcal{Y}$, which is a continuous map. Then the sequence of induced measures $(\mu_n)_{n\in\mathbb{N}}$ given by $\mu_n = \rho_n \mathrm{Pr}_1^{-1}$ satisfies the LDP on $L_\infty([0,T])$ with good rate function

$$\tilde{I}_Q(\phi) = \inf_{(\phi,\eta)\in Pr_1^{-1}(\{\phi\})} I(\phi,\eta) = \inf_{\eta\in L_\infty([0,T])} I(\phi,\eta). \tag{2.106}$$

We can now state the LDP for the input process $(Q_n^+(t))_{t\in[0,T]}$.

Proposition A.7 (LDP for the input process).
The family of probability measures $(\mu_n)_{n\in\mathbb{N}}$ satisfies the LDP on $L_\infty[0,T]$ with rate n and with good rate function $\hat{I}$ given by

$$\hat{I}_Q(\Gamma) = \inf_{\phi\in\Gamma} \tilde{I}_Q(\phi). \tag{2.107}$$

In particular, if Γ is $\hat{I}_Q$-continuous, i.e., $\hat{I}_Q(\Gamma) = \hat{I}_Q(\mathrm{int}(\Gamma)) = \hat{I}_Q(\mathrm{cl}(\Gamma))$, then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Q_n^+([0,T]) \in \Gamma\big) = -\hat{I}_Q(\Gamma). \tag{2.108}$$

Proof. The claim follows from the contraction principle (see [40]). □

It is interesting to look at a specific subset of $L_\infty([0,T])$ that gives good bounds for the input process. We are now in a position to prove Proposition 2.2.1.

Proof. If we compute the Fenchel-Legendre transforms Λ_Q^* and Λ_N^*, and we pick $\eta(t) = \lambda t$ and $\phi(t) = (1/\mu)\eta(t) = (1/\mu)\lambda t$, we can easily check that the rate function attains its minimal value zero. Hence, with high probability as $r \to \infty$, the input process is close to this deterministic path.

We can now estimate the probability of the scaled input process to go outside $\Gamma_{T,\delta}$, which represents a tube of width 2δ around the mean path in the interval $[0,T]$. More precisely,

$$\Gamma_{T,\delta} = \left\{ \gamma \in L_\infty([0,T])\colon\ \frac{\lambda}{\mu}t - \delta < \gamma(t) < \frac{\lambda}{\mu}t + \delta \ \ \forall\, t \in [0,T] \right\}. \tag{2.109}$$

We may set $T = 1$ for simplicity and look at the scaled input process in the time interval $[0,1]$. We have

$$\hat{I}_Q((\Gamma_{1,\delta})^c) = \hat{I}_Q(\mathrm{int}((\Gamma_{1,\delta}))^c) = \hat{I}_Q(\mathrm{cl}((\Gamma_{1,\delta}))^c). \tag{2.110}$$

Hence $(\Gamma_{1,\delta})^c$ is $\hat{I}_Q$-continuous, and so according to (2.108),

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\big(Q_n^+([0,1]) \notin \Gamma_{1,\delta}\big) = -\hat{I}_Q((\Gamma_{1,\delta})^c). \tag{2.111}$$

Since

$$\begin{aligned}
&\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\big(Q_n^+([0,1])\notin\Gamma_{1,\delta}\big)\\
&=\lim_{n\to\infty}\frac{1}{n}\log\mathbb{P}\bigg(\bigg\{\frac{\lambda}{\mu}t-\delta<Q_n^+(t)<\frac{\lambda}{\mu}t+\delta\ \ \forall t\in[0,1]\bigg\}^c\bigg) \qquad (2.112)\\
&=\lim_{S\to\infty}\frac{1}{S}\log\mathbb{P}\bigg(\bigg\{\frac{\lambda}{\mu}s-\delta S<Q^+(s)<\frac{\lambda}{\mu}s+\delta S\ \ \forall s\in[0,S]\bigg\}^c\bigg),
\end{aligned}$$

where we put $s=nt$ and $S=n$, we conclude that the probability to go out of $\Gamma_{S,\delta S}$ is

$$\mathbb{P}\bigg(\bigg\{\frac{\lambda}{\mu}s-\delta S<Q^+(s)<\frac{\lambda}{\mu}s+\delta S\ \ \forall s\in[0,S]\bigg\}^c\bigg)=e^{-S\,\hat{I}_Q((\Gamma_{1,\delta})^c)\,[1+o(1)]},\qquad S\to\infty. \tag{2.113}$$

Because I_Q is convex, to compute $\hat{I}_Q((\Gamma_{1,\delta})^c)$ it suffices to minimise over the linear paths. The minimizer turns out to be one of the two linear paths that go from the origin $(0,0)$ to $(1,\lambda/\mu\pm\delta)$, i.e., $\gamma^*(t)=kt$ with $k=(\lambda\pm\delta\mu)/\mu$. By construction, $\hat{I}_Q((\Gamma_{1,\delta})^c)=\tilde{I}_Q(\gamma^*)=\inf_{\eta\in L_\infty([0,1])}I(\gamma^*,\eta)$, where

$$I(\gamma^*,\eta)=\int_0^1\Lambda_Q^*\left(\frac{d\gamma^*(t)}{d\eta(t)}\right)d\eta(t)+\int_0^1\Lambda_N^*(\dot{\eta}(t))\,dt. \tag{2.114}$$

We want to minimize the sum over all paths η such that $\eta(0)=0$. Both integrals are convex as a function of γ^* and η, hence they are minimized by linear paths. Our choice of $\gamma^*(t)=kt$ is linear, so we set $\eta(t)=ct$ with some constant $c>0$. We can then write

$$\begin{aligned}
I(\gamma^*,\eta)&=\int_0^{\eta(1)}\Lambda_Q^*\left(\frac{d\gamma^*(t)}{cdt}\right)cdt+\int_0^1\Lambda_N^*(\dot{\eta}(t))\,dt\\
&=\int_0^c\Lambda_Q^*\left(\frac{k}{c}\right)cdt+\int_0^1\Lambda_N^*(c)\,dt\\
&=c\left[\frac{k}{c}\mu-1-\log\left(\frac{k\mu}{c}\right)\right]+c\log\left(\frac{c}{\lambda}\right)-c+\lambda.
\end{aligned} \tag{2.115}$$

The value of c that minimizes the right-hand side is $c=\sqrt{\lambda k\mu}$. Substituting this into the formula above, we get

$$K_\delta=\tilde{I}_Q(\gamma^*)=k\mu+\lambda-2\sqrt{\lambda k\mu}=(\lambda+\delta\mu)+\lambda-2\sqrt{\lambda(\lambda+\delta\mu)}. \tag{2.116}$$

Note that $K_\delta>0$ for all $\delta>0$ and $\lim_{\delta\to0}K_\delta=0$. This completes the proof. □

§B Appendix: the output process

The main goal of this appendix is to prove Proposition 2.2.4 in Section 2.2. In Section B.1 we show a lower bound for the output process for the nodes in U, in a setting where the nodes in U are not influenced by the nodes in V. We study the network evolution up to time T_U. In Section B.2 we show that, until the pre-transition time, the network in the internal model behaves actually as we described.

§B.1 The output process in the isolated model

Recall that in the isolated model a node in U keeps activating and deactivating independently of the nodes in V, until its queue length hits zero. We again consider a single queue for a node in U and for simplicity suppress its index. In order to show that, with high probability as $r \to \infty$, the output process $t \mapsto Q^-(t) = cT(t)$ when properly rescaled is close to a deterministic path, we will provide a lower bound for the output process. The upper bound $Q^-(t) \leq ct$ is trivial and holds for any $t \geq 0$, by the definition of output process.

Lemma B.1 (Auxiliary output process).
For all $\delta > 0$ and T large, the following statements hold.

(i) With high probability as $r \to \infty$, the process

$$Q^{\mathrm{LB},T}(t) = \gamma_U r + \rho_U t - \delta T - ct, \qquad t \in [0,T], \tag{2.117}$$

is a lower bound for the actual queue length process $(Q(t))_{t\in[0,T]}$.

(ii) The probability of the lower bound in (i) failing is

$$\frac{1}{2} e^{-K_\delta T\,[1+o(1)]}, \qquad T \to \infty, \tag{2.118}$$

with $K_\delta = (\lambda + \delta\mu) + \lambda - 2\sqrt{\lambda(\lambda+\delta\mu)}$.

Proof. We prove the two statements separately.

(i) By Proposition 2.2.1, with high probability as $r \to \infty$, we have $Q^+(t) \geq \rho_U t - \delta T$ for any $\delta > 0$. Trivially, $Q^-(t) \leq ct$. It is therefore immediate that, with high probability as $r \to \infty$, $Q^{\mathrm{LB},T}(t) \leq Q(t)$.

(ii) The exponentially small probability of $Q^+(t)$ going below the lower bound is half of the probability given by Proposition 2.2.1, i.e.,

$$\frac{1}{2} e^{-K_\delta T\,[1+o(1)]}, \qquad T \to \infty, \tag{2.119}$$

with $K_\delta = (\lambda + \delta\mu) + \lambda - 2\sqrt{\lambda(\lambda+\delta\mu)}$.

□

We study the network evolution up to time T_U defined in Definition 2.2.2, the expected time a single node queue takes to hit zero. We will prove in Appendix B.2 that, with high probability as $r \to \infty$, the pre-transition time in the internal model coincides in distribution with the pre-transition time in the isolated model, which occurs before T_U. Hence it is enough to study the isolated model up to T_U.

Definition B.2 (Auxiliary times).
We next define two times that will be useful in our analysis.

(T_U^*) Consider the auxiliary output process $Q^{\mathrm{LB},T_U}(t)$ up to time T_U. We define T_U^* as the time needed for the process to hit zero, i.e.,

$$T_U^* = T_U^*(r) = \frac{\gamma_u r - \delta T_U}{c - \rho_U} = \frac{\gamma_u - \delta\alpha}{c - \rho_U} r = \alpha' r, \tag{2.120}$$

with $\alpha' = \frac{\gamma_u - \delta\alpha}{c - \rho_U}$. The difference $T_U - T_U^* = \frac{\delta\alpha}{c-\rho_U} r$ is of order r. The queue length at time T_U^* is not zero, but still of order r.

(T_U^{**}) We define a smaller time T_U^{**} in such a way that, not only $Q(T_U^{**}) \asymp r$, but also $Q^{\mathrm{LB},T_U}(T_U^{**}) \asymp r$, i.e.,

$$T_U^{**} = T_U^{**}(r) = T_U - 2(T_U - T_U^*) = \left(\frac{\gamma_U - 2\delta\alpha}{c - \rho_U}\right) r = \alpha'' r, \tag{2.121}$$

with $\alpha'' = \frac{\gamma_U - 2\delta\alpha}{c - \rho_U}$.

Definition B.3 (Inactivity process).
Define the *inactivity process* by setting

$$W(t) = t - T(t), \tag{2.122}$$

which equals the total amount of inactivity time until time t.

Recall that the service process $t \mapsto Q^-(t)$ with $Q^-(0) = 0$ is an alternating sequence of activity periods and inactivity periods. The activity periods Z_i, $i \in \mathbb{N}$, are i.i.d. exponential random variables with mean 1. The inactivity periods W_m, $m \in \mathbb{N}$, are exponential random variables with a mean that depends on the actual queue length at the time when each of these periods starts, namely, if $W_m = [t_m^{(i)}, t_m^{(f)}]$, then $W_m \simeq \mathrm{Exp}(g_U(Q(t_m^{(i)})) + O(1/r))$. The queue length during this inactivity intervals is actually increasing, but we are considering very small intervals, whose lengths are of order $1/r$, so that the queue length does not change much and the error is then $O(1/r)$.

To state our lower bound on the output process, we need the following two lemmas.

Lemma B.4 (Upper bound on number of activity periods).
Let $M(t)$ be the number of activity periods that end before time t. Then, for all $\epsilon_1 > 0$, the following statements hold.

(i) With high probability as $r \to \infty$,

$$M(T_U^{**}) \le (1 + \epsilon_1) T_U^{**}. \tag{2.123}$$

(ii) The probability of the upper bound in (i) failing is

$$\mathbb{P}_u(M(T_U^{**}) > (1+\epsilon_1)T_U^{**}) \le e^{-K_1 r\,[1+o(1)]} + \frac{1}{2} e^{-K_\delta \alpha r\,[1+o(1)]}, \qquad r \to \infty, \tag{2.124}$$

with $K_1 = \alpha'' \frac{\epsilon_1 - \log(1+\epsilon_1)}{1+\epsilon_1}$, K_δ as in Lemma B.1

Proof. We prove the two statements separately.

(i) Note that $M(T_U^{**})$ counts the number of activity periods before time T_U^{**}, each of which has an average duration 1. Since activity periods alternate with inactivity periods, we expect $M(T_U^{**})$ to be less than T_U^{**}. Assume now, for small $\epsilon_1 > 0$, that $M(T_U^{**}) > (1+\epsilon_1)T_U^{**}$, which means that the number of activity periods before T_U^{**} is greater than the length of the interval $[0, T_U^{**}]$. This implies that the average length of each activity period before time T_U^{**} is strictly less than 1, namely, that $\frac{1}{T_U^{**}} \sum_{i=1}^{T_U^{**}} Z_i \leq 1/(1+\epsilon_1)$. According to Cramér's theorem, we can compute the probability of this last event as

$$\mathbb{P}_u\left(\sum_{i=1}^{T_U^{**}} Z_i \leq \left(\frac{1}{1+\epsilon_1}\right) T_U^{**}\right) = e^{-T_U^{**} I\left(\frac{1}{1+\epsilon_1}\right)[1+o(1)]}, \qquad r \to \infty, \tag{2.125}$$

with rate function $I(x) = x\log(x) - x + 1$. Therefore, it occurs with exponentially small probability. Hence $M(T_U^{**}) > (1+\epsilon_1)T_U^{**}$ must also occur with a probability which is also exponentially small. With high probability as $r \to \infty$, we then have that

$$M(T_U^{**}) \leq (1+\epsilon_1)T_U^{**}. \tag{2.126}$$

Recall that $T_U^{**} = \alpha'' r$. The counting of alternating activity and inactivity periods gets affected when the queue length hits zero, since then the node deactivates and the lengths of the activity periods are not regular anymore. At time T_U^{**}, with high probability as $r \to \infty$, the queue length is still of order r. Hence, the probability that it hits zero at any time in the interval $[0, T_U^{**}]$ is very small, since this event would imply the node to have a queue length that is below the lower bound, $Q(T_U^{**}) \leq Q^{\mathrm{LB},T_U}(T_U^{**})$, which happens with an exponentially small probability by Lemma B.1.

(ii) We can write

$$\begin{aligned}\mathbb{P}_u(M(T_U^{**}) > (1+\epsilon_1)T_U^{**}) &\leq e^{-T_U^{**} I\left(\frac{1}{1+\epsilon_1}\right)[1+o(1)]} + \frac{1}{2} e^{-K_\delta T_U [1+o(1)]} \\ &= e^{-K_1 r [1+o(1)]} + \frac{1}{2} e^{-K_\delta \alpha r [1+o(1)]}, \qquad r \to \infty,\end{aligned} \tag{2.127}$$

with $K_1 = \alpha'' I\left(\frac{1}{1+\epsilon_1}\right) = \alpha'' \frac{\epsilon_1 - \log(1+\epsilon_1)}{1+\epsilon_1}$, K_δ as in Lemma B.1.

□

Lemma B.5 (Upper bound on inactivity process).
For all $\delta, \epsilon_1, \epsilon_2 > 0$ small, the following statements hold.

(i) With high probability as $r \to \infty$,

$$W(T_U^{**}) \leq \epsilon_2 r. \tag{2.128}$$

(ii) The probability of the upper bound in (i) failing is

$$\mathbb{P}_u\big(W(T_U^{**}) > \epsilon_2 r\big) \le e^{-K_\delta \alpha r\,[1+o(1)]} + e^{-K_1 r\,[1+o(1)]} \\ + e^{-\left(K_2 r + K_3 \frac{r}{g_U(r)} + K_4 r \log g_U(r)\right)[1+o(1)]}, \qquad r \to \infty, \tag{2.129}$$

with $K_2 = \alpha''(1+\epsilon_1)\big(-1-\log\big(\frac{\epsilon_2}{\alpha''(1+\epsilon_1)}\big)\big), K_3 = \epsilon_2, K_4 = \alpha''(1+\epsilon_1)$.

Proof. We prove the two statements separately.

(i) Since $M(t)$ counts the number of activity periods, and we start with an active node (initially all nodes in U are active), we have

$$W(T_U^{**}) \le \sum_{m=1}^{M(T_U^{**})} W_m \le \sum_{m=1}^{M(T_U^{**})} \hat{W}_m, \tag{2.130}$$

where $\hat{W}_m$ are i.i.d. exponential random variables with rate $g_U(Q^{\mathrm{LB},T_U}(T_U^{***}))$, and T_U^{***} is the starting point of the last inactivity period before time T_U^{**}. By the construction of T_U^{**}, we know that $Q^{\mathrm{LB},T_U}(T_U^{***})$ is of order r. The last inactivity period is expected to be longer than the previous ones, since the activation rates depend on the actual queue length, which is decreasing over time. To make the inactivity periods $\hat{W}_m$ longer, we consider the lower bound $Q^{\mathrm{LB},T_U}(t)$ for the actual queue length given in Lemma B.1.

By Lemma B.4, with high probability as $r \to \infty$, $M(T_U^{**}) \le (1+\epsilon_1)T_U^{**}$, and so

$$W(T_U^{**}) \le \sum_{m=1}^{M(T_U^{**})} \hat{W}_m \le \sum_{m=1}^{(1+\epsilon_1)T_U^{**}} \hat{W}_m. \tag{2.131}$$

Define $n = [(1+\epsilon_1)T_U^{**}]$. By Cramér's theorem, for small $\epsilon_3 > 0$,

$$\mathbb{P}_u\left(\sum_{m=1}^{(1+\epsilon_1)T_U^{**}} \hat{W}_m \ge \epsilon_3 T_U^{**}\right) \le \mathbb{P}_u\left(\sum_{m=1}^{n} \hat{W}_m \ge \frac{\epsilon_3}{1+\epsilon_1} n\right) \\ = e^{-nI\left(\frac{\epsilon_3}{1+\epsilon_1}\right)[1+o(1)]} \\ = e^{-T_U^{**}(1+\epsilon_1)I\left(\frac{\epsilon_3}{1+\epsilon_1}\right)[1+o(1)]}, \qquad n \to \infty, \tag{2.132}$$

where I is the rate function given by

$$I(x) = \frac{x}{g_U\big(Q^{\mathrm{LB},T_U}(T_U^{***})\big)} - 1 - \log x + \log g_U\big(Q^{\mathrm{LB},T_U}(T_U^{***})\big). \tag{2.133}$$

We take $\epsilon_3 > (1+\epsilon_1)/g_U\big(Q^{\mathrm{LB},T_U}(T_U^{***})\big) \asymp 1/g_U(r)$ arbitrarily small, so that we can apply Cramér's theorem. Combining (2.131)–(2.132), we obtain that, with high probability as $r \to \infty$,

$$W(T_U^{**}) \le \epsilon_3 T_U^{**} = \epsilon_3 \alpha'' r = \epsilon_2 r, \tag{2.134}$$

where $\epsilon_2 = \epsilon_3 \alpha''$ can be taken arbitrarily small.

(ii) We can write

$$\begin{aligned}
&\mathbb{P}_u\left(\sum_{m=1}^{(1+\epsilon_1)T_U^{**}} \hat{W}_m > \epsilon_3 T_U^{**}\right)\\
&\leq e^{-T_U^{**}(1+\epsilon_1)I\left(\frac{\epsilon_3}{1+\epsilon_1}\right)[1+o(1)]}\\
&= e^{-\alpha'' r(1+\epsilon_1)\left(\frac{\epsilon_3}{(1+\epsilon_1)g_U(r)}-1-\log\left(\frac{\epsilon_3}{1+\epsilon_1}\right)+\log g_U(r)\right)[1+o(1)]}\\
&= e^{-\left[\alpha''(1+\epsilon_1)\left(-1-\log\left(\frac{\epsilon_3}{1+\epsilon_1}\right)\right)r+\epsilon_3\alpha''\frac{r}{g_U(r)}+\alpha''(1+\epsilon_1)r\log(g_U(r))\right][1+o(1)]}\\
&= e^{-\left(K_2 r+K_3\frac{r}{g_U(r)}+K_4 r\log g_U(r)\right)[1+o(1)]}, \qquad r\to\infty,
\end{aligned} \tag{2.135}$$

where the constants $K_2 = \alpha''(1+\epsilon_1)\big(-1-\log\big(\frac{\epsilon_2}{\alpha''(1+\epsilon_1)}\big)\big)$, $K_3 = \epsilon_3\alpha'' = \epsilon_2$ and $K_4 = \alpha''(1+\epsilon_1)$. We also have to consider the probabilities computed in (2.118) and (2.124), and the claim in (2.129) is settled.

□

We are now in a position to prove Proposition 2.2.4.

Proof. The equation $Q^-(t) \geq ct - \epsilon r$ can be read as $T(t) \geq t - \epsilon r/c$. This is equivalent to saying that $W(t) \leq \epsilon r/c$ for all $t \in [0, T_U]$. By taking $\epsilon_2 = \epsilon/(3c)$ in Lemma B.5, we know that, for all $t \in [0, T_U^{**}]$, $W(t) \leq W(T_U^{**}) \leq \epsilon r/(3c)$. Moreover, in the interval $[T_U^{**}, T_U]$, the cumulative amount of inactivity time is trivially bounded from above by the length of the interval, which is $\frac{2\delta r}{c-\rho_U} \leq 2\epsilon r/(3c)$, and ϵ can be taken arbitrarily small, since δ can be taken arbitrarily small. Putting the two bounds together, we find that, with high probability as $r \to \infty$,

$$W(t) \leq \epsilon_2 r + \frac{2\delta r}{c-\rho_U} \leq \frac{1}{3}\frac{\epsilon r}{c} + \frac{2}{3}\frac{\epsilon r}{c} = \frac{\epsilon r}{c}, \qquad t \in [0, T_U], \tag{2.136}$$

and the probability of this not happening is given by (2.129). □

The above lower bound $Q^-(t) \geq ct - \epsilon r$ and the trivial upper bound $Q^-(t) \leq ct$ imply that, with high probability as $r \to \infty$, the output process $Q^-(t)$ stays close to the path $c \mapsto ct$ by sending ϵ to zero. In other words, the node stays almost always active all the time before T_U.

§B.2 The output process in the internal model

In this section we want to couple the isolated model and the internal model and show that they have identical behavior in the time interval $[0, \tau_G^{\mathrm{int}}]$. Hence it follows that the output process in the internal model for nodes in U actually behaves as in the isolated model described in Section B.1, until the pre-transition time.

Proposition B.6 (Coupling the internal and the isolated model).
Let $X_i^{\mathrm{int}}(t)$ and $X_i^{\mathrm{iso}}(t)$ denote the activity state of a node i at time t in the internal and the isolated model, respectively. Then

$$\lim_{r\to\infty} \mathbb{P}_u\big(X_i^{\mathrm{int}}(t) = X_i^{\mathrm{iso}}(t)\ \ \forall\, i \in U \sqcup V\ \forall\, t \in [0, \tau_G^{\mathrm{int}}]\big) = 1. \tag{2.137}$$

Consequently, with high probability as $r \to \infty$, the pre-transition times in the internal and the isolated model coincide, i.e.,

$$\lim_{r\to\infty} \mathbb{P}_u(\tau_G^{\text{int}} = \tau_G^{\text{iso}}) = 1. \tag{2.138}$$

Proof. In Section B.1 we determined upper and lower bounds for the output process for nodes in U in the isolated model up to time T_U. Assume now that $\tau_G^{\text{int}} \leq T_U$. When considering the internal model and the set of nodes in V, note that these bounds are not true for the whole interval $[0, T_U]$, since at time τ_G^{int} some nodes in V already start to activate and influence the behavior of nodes in U.

If we look at the interval $[0, \tau_G^{\text{int}}]$, then we note that the queue length process for a node $i \in U$ is not affected by nodes in V, and so it behaves in exactly the same way as if the node were isolated. The activation and deactivation Poisson clocks at node i are synchronized, and are ticking at the same time in the isolated model and in the internal model, so that $X_i^{\text{int}}(t) = X_i^{\text{iso}}(t)$. Moreover, the activity states of nodes in V are always equal to 0 in both models. Hence we conclude that the activity states of every node coincide up to the pre-transition time τ_G^{int}. Consequently, the pre-transition times in the internal and the isolated model coincide on the event $\{\tau_G^{\text{int}} \leq T_U\}$, which can then be written as the event $\{\tau_G^{\text{iso}} \leq T_U\}$. For the latter we know that it has a high probability as $r \to \infty$ (see proof of Proposition 2.3.8 in Section 2.3.2). □