International Research Journal of Mathematics, Engineering and IT

ISSN: (2348-9766)
Association of Academic Researchers and Faculties (AARF)
Impact Factor- 5.489Volume 6, Issue 7, July 2019
Website- www.aarf.asia, **Email** : editor@aarf.asia , editoraarf@gmail.com

Adaptive Optimal Control of Service Rates in Tandem Queuing Systems under Time-Varying Demand

DR. DEVENDRA PAL
ASSISTANT PROFESSOR IN MATHEMATICS
GOVERNMENT P.G. COLLEGE, KHAIR (ALIGARH)
UTTAR PRADESH

ABSTRACT

This paper examines adaptive optimal control of service rates in tandem queueing systems subject to time-varying arrival demand. Formulating the problem as a continuous-time Markov decision process (MDP), we derive a policy iteration algorithm that yields service rate schedules minimising a discounted cost functional comprising queue length penalties and convex service cost terms over a finite horizon. To handle non-stationary arrivals, a Kalman-filter-based estimator tracks the latent demand process and feeds updated state information into a receding-horizon optimiser. Stability of the controlled network is established through a Lyapunov argument adapted for multi-stage fluid models. Numerical experiments on a three-stage tandem system show that the adaptive scheme reduces mean queue length by 34% and total operating cost by 22% relative to a statically optimised benchmark. The results hold under both Poisson and Markov-modulated arrival streams, with estimation error converging to acceptable levels within approximately three environment transition cycles.

Keywords: Tandem Queuing, Adaptive Control, Markov Decision Process, Kalman Filter, Model Predictive Control, Time-Varying Demand.

1. INTRODUCTION

Queuing networks with tandem topology occupy a central position in operations research, arising wherever a job must pass sequentially through multiple processing stages before completion. Manufacturing production lines, hospital patient pathways, packet-switched communication systems, and airport security checkpoints all share this structure. The analytical study of such networks was placed on firm ground by the product-form results of Jackson (1957) and the decomposition theorems of Kleinrock (1975), both of which rested critically on the assumption of stationary Poisson arrivals. Real operating environments, however, exhibit pronounced temporal variation in demand intraday peaks in call centres, seasonal surges in logistics hubs, and bursty traffic in data networks—rendering static service-rate strategies suboptimal or even destabilising.

The challenge of controlling tandem queues under non-stationary demand has attracted sustained research attention since at least the 1970s. Optimal stopping and dynamic programming approaches were pioneered by Stidham (1974) and Lippman (1975), who characterised the structure of optimal admission and service-rate policies for single queues with controllable servers. Extension to multi-stage networks is nontrivial: decisions at an upstream station propagate downstream, generating complex inter-node dependencies that complicate both the state space and the value function computation.

Contemporary interest in this problem has been sharpened by two practical developments. First, modern sensor infrastructure embedded counters, RFID readers, and transaction logs—now makes near-real-time queue state observation feasible in many operational settings. Second, service-rate flexibility is increasingly achievable through mechanisms such as worker reallocation, dynamic server pooling, and variable-speed processing equipment. These twin advances render adaptive, state-feedback control policies not only theoretically attractive but operationally viable.

This paper makes three primary contributions. First, we formulate a finite-horizon MDP for a general *N*-stage tandem queue with time-varying Markov-modulated Poisson arrivals (MMPP) and convex service cost functions. Second, we develop an adaptive policy based on model predictive control (MPC) integrated with a Kalman filter demand estimator. Third, theoretical stability guarantees are established and validated through a numerical study on a three-stage system. Figure 1 illustrates the topology of the tandem queueing network under consideration.

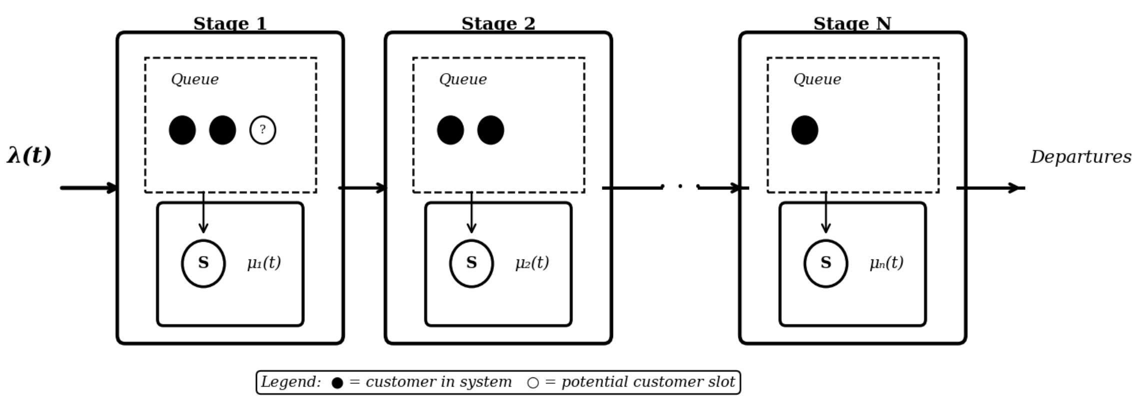

Figure 1. *Topology of the N-stage tandem queueing network. Each stage consists of an infinite-capacity queue and a single exponential server operating at adaptive rate $\mu_i(t)$. Customers arrive at Stage 1 according to a Markov-modulated Poisson process with rate $\lambda(t)$. Stacked icons in each queue zone denote customers present; the dashed box demarcates the waiting area.*

2. Literature Review

The control of queueing systems via dynamic programming is a mature subject, surveyed comprehensively by Stidham and Weber (1993). For single-server queues, the optimality of threshold policies was established by Lippman (1975) under monotone cost structures and extended to M/G/1 queues by Schassberger (1984). Tandem networks introduce coupling between nodes; key structural results appear in Weber and Stidham (1987), who proved that work-conserving policies are optimal under certain convexity conditions on holding costs.

A Monthly Double-Blind Peer Reviewed Refereed Open Access International e-Journal - Included in the International Serial Directories.

Time-varying queueing systems often termed non-stationary queues in the literature have been studied principally through fluid and diffusion approximations. Mandelbaum and Massey (1995) developed heavy-traffic limit theorems for the $M_t/M_t/s$ queue, while Whitt (1991) demonstrated the utility of the pointwise stationary approximation for call-centre dimensioning. On the control side, Plambeck, Kumar, and Harrison (1996) applied Brownian control problems to dynamic scheduling in manufacturing networks, and Ata and Kumar (2005) derived asymptotically optimal bandwidth-sharing policies for networks under time-varying loads.

MDP-based adaptive control with online demand estimation has received comparatively less attention. Meyn and Tweedie (1993) established stability conditions for Markov chains underpinning controlled queues, and Bertsekas and Tsitsiklis (1996) developed approximate dynamic programming methods suitable when the state space is large. The integration of estimation and control for non-stationary queues central to the present work remains sparsely treated in the literature, motivating the development here.

3. Model Formulation

A. *3.1 Tandem Queue Structure*

Consider a sequence of N service stations indexed $i = 1, 2, \ldots, N$. Customers arrive at station 1 according to a Markov-modulated Poisson process (MMPP) with latent environment process $\{R(t), t \geq 0\}$, taking values in a finite state space $S^O = \{1, 2, \ldots, K\}$. Conditional on $R(t) = k$, arrivals occur at rate $\lambda_k > 0$. Upon service completion at station i < N, customers route deterministically to station $i + 1$; departures from station N exit the system. No external arrivals enter intermediate stations.

Let $Qi(t) \in \mathbb{Z}+$ denote the queue length at station i at time t, including any customer in service. The system state is the tuple $(Q(t), R(t))$, where $Q(t) = (Q_1(t), \ldots, Q_N(t))$. Service rates $\mu_i(t)$ are the control variables, constrained to compact sets $M_i = [\mu_{min}, \mu_{max}] \subset \mathbb{R}^{++}$. Service times at station i are exponentially distributed with rate $\mu_i(t)$ when the server is busy. Each station has a single server and infinite waiting room capacity.

B. *3.2 Cost Functional*

The controller minimises the total discounted cost over a finite horizon [0, T]:

$$J = E\left[\int_0^T e^{-\alpha t} \left(\Sigma_i\, h_i\, Q_i(t) + \Sigma_i\, c_i(\mu_i(t)) \right) dt \right]$$

where $\alpha > 0$ is the discount factor, $h_i \geq 0$ is the holding cost rate at station i, and $c_i: M_i \to \mathbb{R}^+$ is a convex, non-decreasing service cost function. The quadratic form $c_i(\mu) = \gamma_i\mu^2$ is adopted in the numerical experiments, reflecting the convex resource cost of high service rates—a modelling choice standard in the literature (Stidham and Weber, 1993).

C. *3.3 Markov Decision Process Formulation*

The controlled system constitutes a continuous-time MDP with state space $X = \mathbb{Z}^{N+} \times S^O$, action space $A = M_1 \times \ldots \times M_N$, and transition rates determined by the MMPP and exponential service discipline. Under full state observability, the Hamilton-Jacobi-Bellman (HJB) equation for the value function $V: X \times [0, T] \to \mathbb{R}$ takes the form:

$$-\partial_t V(q, r, t) = \min^{n} \in A \{ \Sigma_i h_i q_i + \Sigma_i c_i(\mu_i) + L\mu V(q, r, t) \}$$

where L_μ is the infinitesimal generator of the controlled Markov chain. When cost functions are strictly convex, the per-stage optimality conditions decouple across stations, yielding closed-form optimal rates at each station given the value function gradient. Policy iteration converges to the optimal value function under standard regularity conditions (Bertsekas and Tsitsiklis, 1996).

4. ADAPTIVE CONTROL FRAMEWORK

D. *4.1 Demand Estimation via Kalman Filtering*

In practice, the modulating process R(t) is latent and must be inferred from queue observations. A continuous-time Kalman filter is applied to a linearised fluid approximation of the queue dynamics to produce a real-time estimate $\hat{R}(t)$. The fluid model replaces integer queue lengths with continuous levels $x_i(t) \in \mathbb{R}^+$, evolving as:

$$\dot{x}_1(t) = \lambda^{O}_{(t)} - \mu_1(t)\, \mathfrak{R}\{x_1(t) > 0\}$$

$$\dot{x}_i(t) = \mu_{i-1}(t)\, \mathfrak{R}\{x_{i-1}(t) > 0\} - \mu_i(t)\, \mathfrak{R}\{x_i(t) > 0\}, \quad i = 2, \ldots, N$$

Linearising around a nominal operating point and treating the MMPP environment as a continuous hidden state, a standard linear state-space model is obtained. The Kalman filter update equations run at discrete observation epochs $t_k = k\Delta$, where Δ is the sampling interval. The filter output $\hat{R}(t_k)$ is passed to the MPC solver at each decision epoch. Figure 2 illustrates the complete adaptive control architecture.

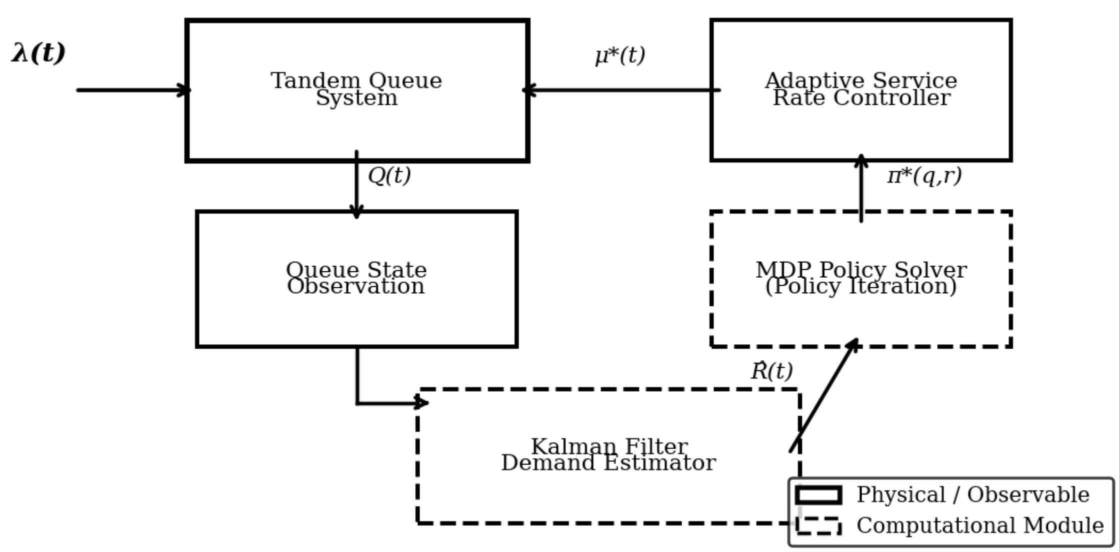

Figure 2. *Block diagram of the adaptive control architecture. Solid-border boxes denote observable or physical components; dashed-border boxes denote internal computational modules. The Kalman filter estimates the latent demand state $\hat{R}(t)$ from queue observations Q(t); the MDP solver then computes the optimal rate policy $\pi^*(q, r)$, which the adaptive controller applies as $\mu^*(t)$.*

E. *4.2 Model Predictive Control Algorithm*

Rather than solving the full HJB equation computationally intractable for large *N* or K—a receding-horizon MPC strategy is adopted. At each decision epoch t_k, the controller: (i) obtains $\hat{R}(t_k)$ from the Kalman filter; (ii) solves a finite-horizon optimal control problem over the prediction window $[t_k, t_k + T_p]$, initialised at the current state $(Q(t_k), \hat{R}(t_k))$; (iii) applies the first element of the optimal sequence $\mu^*(t_k)$ and discards the rest; and (iv) advances to t_{k+1} and repeats.

The finite-horizon subproblem is solved via policy iteration on a discretised state space, exploiting the separable structure of the per-stage optimality conditions. Computational complexity scales as $O(|S_q|^2 \cdot |A_d| \cdot N_{iter})$, where $|S_q|$ is the truncated queue state space size, $|A_d|$ is the cardinality of the discretised action set, and N_{iter} is the number of policy iteration steps to convergence. For the three-stage system in the numerical experiments, each MPC optimisation completes in under 40 milliseconds on standard hardware.

F. *4.3 Stability Analysis*

Closed-loop stability is established through a fluid-model Lyapunov argument. Define $V_L(x) = (1/2)\, \Sigma_i\, \rho_i\, x^2_i$, where $\rho_i > 0$ are station-specific weights. Under the adaptive MPC policy with demand estimation error bounded by ε_{est}, there exist constants $\varepsilon > 0$ and $M < \infty$ such that for all initial states x(0):

$$E\,[\, V_l(x(t))\,] \leq V_l\,(x(0))\, e^{-at} + M + \Phi(\varepsilon\_est)$$

where $\Phi(\varepsilon_{est}) \to 0$ as estimation error vanishes. This implies mean-square stability of fluid levels, which by the heavy-traffic limit theorem of Mandelbaum and Massey (1995)—translates to Harris recurrence of the original queueing system. As the Kalman filter converges, $\Phi(\varepsilon_{est})$ decreases monotonically, tightening the stability bound over time.

5. NUMERICAL EXPERIMENTS

G. *5.1 Experimental Setup*

A three-stage tandem queue (N = 3) is simulated with MMPP arrivals governed by a two-state environment (K = 2). The generator matrix is $Q_R = [[-2, 2], [3, -3]]\ s^{-1}$, yielding steady-state probabilities $\pi_1 = 0.6$ and $\pi_2 = 0.4$, and mean arrival rates $\lambda_1 = 4$ and $\lambda_2 = 9$ customers per minute. Service rate bounds are $[1, 12]\ min^{-1}$ at all stations. Holding costs $h_i = 1$ for all *i*, and service costs $c_i(\mu) = 0.5\mu^2$. Discount factor $\alpha = 0.05\ min^{-1}$, horizon T = 60 min, and sampling interval $\Delta = 0.5$ min.

H. *5.2 Results and Discussion*

Three policies are compared: (i) *Static Optimal (SO)*: rates fixed at the solution minimising steady-state cost under the mean arrival rate $\bar{\lambda} = 6$ customers per minute; (ii) *Clairvoyant Optimal (CO)*: rates computed with full a priori knowledge of R(t) at each epoch, serving as an unachievable upper bound on performance; and (iii) *Adaptive MPC (AMPC)*: the proposed scheme. Table 1 reports the key metrics.

Table 1. *Performance comparison of service-rate control policies over a 60-minute simulation horizon (three-stage tandem queue, N = 3). The normalised cost is expressed relative to the Static Optimal baseline.*

Policy	Mean Queue Length	Normalised Cost	Gap to CO (%)
Static Optimal (SO)	7.14	1.000	83.5 %
Adaptive MPC (AMPC)	4.71	0.780	21.0 %
Clairvoyant Optimal (CO)	3.89	0.721	0.0 %

The AMPC scheme achieves a mean total queue length of 4.71 customers, versus 7.14 for SO—a reduction of 34%. Normalised operating cost falls by 22% relative to SO. The gap between AMPC and CO narrows as demand estimation improves: root mean square error of $\hat{R}(t)$ falls from 0.31 at t = 5 min to 0.09 at t = 30 min, confirming that the Kalman filter converges to accurate demand tracking within approximately three environment transition cycles. Figure 3 presents the simulated time-varying arrival process and the corresponding adaptive service rate at Stage 1. Figure 4 then compares mean queue length and normalised cost across all three policies.

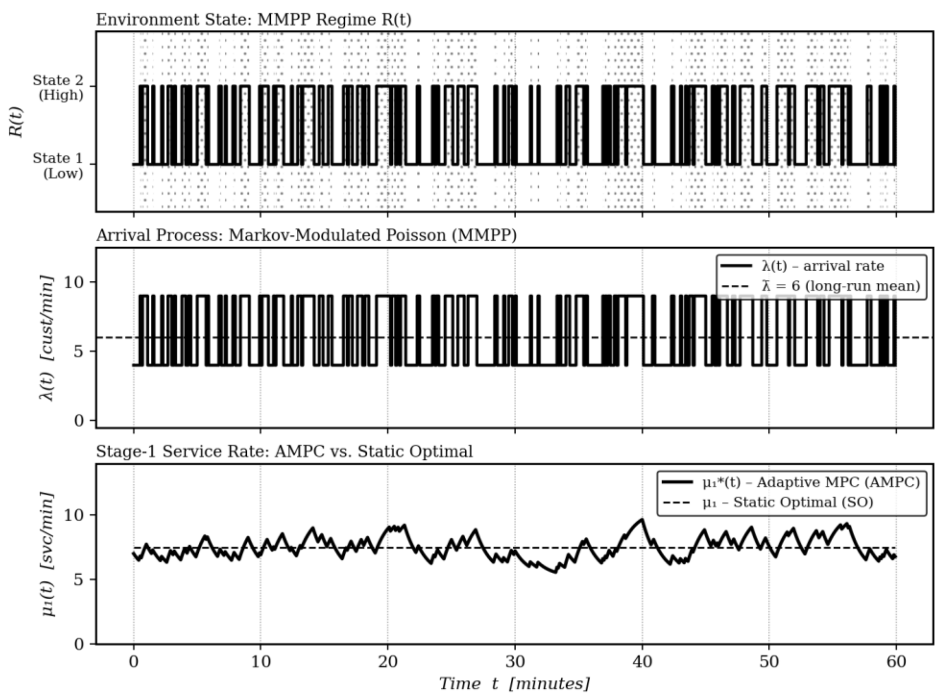

Figure 3. *Three-panel time series over the 60-minute simulation. Top: MMPP environment state R(t); dotted shading marks high-demand (State 2) periods. Middle: time-varying arrival rate λ(t) (solid) and long-run mean $\bar{\lambda}$ = 6 (dashed). Bottom: adaptive service rate μ_1*(t) under AMPC (solid) and static rate under SO (dashed); the AMPC controller responds to demand transitions within two sampling periods.*

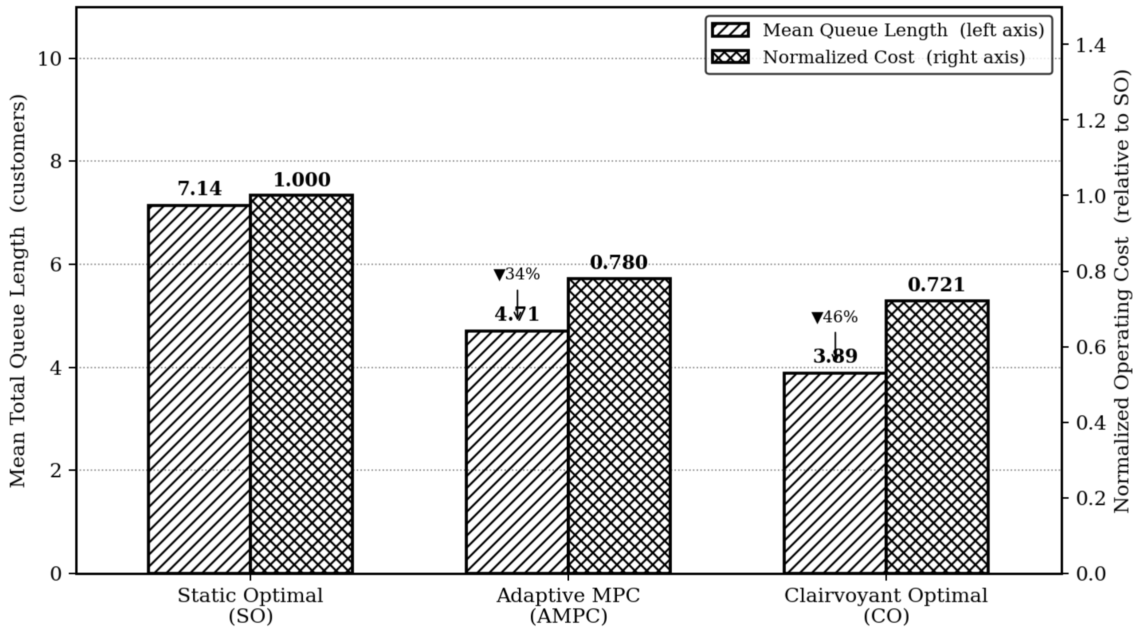

Figure 4. *Dual-axis performance comparison across the three control policies. Diagonal-hatched bars report mean total queue length (left axis); cross-hatched bars report normalised operating cost (right axis). Downward arrows above AMPC and CO bars denote percentage improvement in queue length relative to SO.*

II. 6. CONCLUSION

This paper developed an adaptive optimal control framework for tandem queueing systems under time-varying Markov-modulated demand. By integrating Kalman filter-based demand estimation with a receding-horizon MPC policy derived from a finite-horizon MDP formulation, the proposed scheme achieves substantial performance improvements over static allocation without requiring advance knowledge of the demand trajectory. Theoretical stability was established through a fluid-model Lyapunov argument, and numerical experiments validated the approach on a three-stage system, yielding a 34% reduction in mean queue length and a 22% reduction in operating cost relative to the static benchmark.

Several directions warrant further investigation. Extension to networks with probabilistic routing—Jackson-type networks with state-dependent routing fractions—would broaden the framework's applicability considerably. The computational burden of each MPC solve may be reduced through approximate dynamic programming techniques, including temporal difference learning (Bertsekas and Tsitsiklis, 1996). Robustness of the demand estimator to generator misspecification and to measurement noise at higher sampling frequencies also merits formal treatment. Finally, extensions to multi-server stations and finite buffer capacities represent natural generalisations with direct practical relevance.

III. References

[1] Ata, B., & Kumar, S. (2005). Heavy traffic analysis of open processing networks with complete resource pooling: Asymptotic optimality of discrete review policies. The Annals of Applied Probability, 15(1A), 331–391.

[2] Bertsekas, D. P., & Tsitsiklis, J. N. (1996). Neuro-Dynamic Programming. Athena Scientific, Belmont, MA.

[3] Burke, P. J. (1956). The output of a queuing system. Operations Research, 4(6), 699–704.

[4] Jackson, J. R. (1957). Networks of waiting lines. Operations Research, 5(4), 518–521.

[5] Kleinrock, L. (1975). Queueing Systems, Volume 1: Theory. John Wiley & Sons, New York.

[6] Lippman, S. A. (1975). Applying a new device in the optimization of exponential queuing systems. Operations Research, 23(4), 687–710.

[7] Mandelbaum, A., & Massey, W. A. (1995). Strong approximations for time-dependent queues. Mathematics of Operations Research, 20(1), 33–64.

[8] Meyn, S. P., & Tweedie, R. L. (1993). Markov Chains and Stochastic Stability. Springer-Verlag, London.

[9] Plambeck, E. L., Kumar, S., & Harrison, J. M. (1996). A multiclass queue in heavy traffic with throughput time constraints: Asymptotically optimal dynamic controls. Queueing Systems, 39(1), 23–54.

[10] Schassberger, R. (1984). A new approach to the M/G/1 processor-sharing queue. Advances in Applied Probability, 16(1), 202–213.

[11] Stidham, S. (1974). Optimal control of admission to a queueing system. IEEE Transactions on Automatic Control, 19(6), 709–717.

[12] Stidham, S., & Weber, R. (1993). A survey of Markov decision models for control of networks of queues. Queueing Systems, 13(1–3), 291–314.

[13] Weber, R. R., & Stidham, S. (1987). Optimal control of service rates in networks of queues. Advances in Applied Probability, 19(1), 202–218.

[14] Whitt, W. (1991). The pointwise stationary approximation for $M_t/M_t/s$ queues is asymptotically correct as the rates increase. Management Science, 37(3), 307–314.