A Get out of Residuation Jail Free Card

Minimum Arrivals Bounds in Stochastic Network Calculus

distributed computer systems

Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

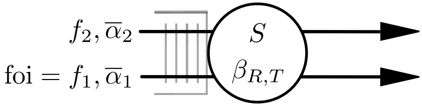
Vlad-Cristian Constantin, Paul Nikolaus, Jens B. Schmitt

June 4th 2025

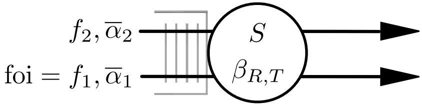
Outline

Motivation

Residuation with a (Min-Plus) Service Curve [BBLC18]



Residuation with a (Min-Plus) Service Curve [BBLC18]



Theorem

Let $t \geq 0$. Consider a system $\mathcal{S}$ that arbitrarily multiplexes two flows f_1 and f_2, where the arrivals of f_2, A_2, are constrained by α_2. Further, assume that $\mathcal{S}$ guarantees a **min-plus** *service curve β to the aggregate of the flows. Then, the leftover service*

$$\xi(t) := \beta(t) - \alpha_2(t)$$

is a service curve for flow f_1.

Residuation with a (Min-Plus) Service Curve [BBLC18]

Theorem

Let $t \geq 0$. Consider a system $\mathcal{S}$ that arbitrarily multiplexes two flows f_1 and f_2, where the arrivals of f_2, A_2, are constrained by α_2. Further, assume that $\mathcal{S}$ guarantees a **min-plus** *service curve β to the aggregate of the flows. Then, the leftover service*

$$\xi(t) := \beta(t) - \alpha_2(t)$$

is a service curve for flow f_1.

- The residual service curve can potentially be negative

Residuation with a (Min-Plus) Service Curve [BBLC18]

Theorem

Let $t \geq 0$. Consider a system $\mathcal{S}$ that arbitrarily multiplexes two flows f_1 and f_2, where the arrivals of f_2, A_2, are constrained by α_2. Further, assume that $\mathcal{S}$ guarantees a **min-plus** *service curve β to the aggregate of the flows. Then, the leftover service*

$$\xi(t) := \beta(t) - \alpha_2(t)$$

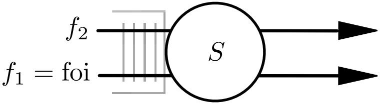
is a service curve for flow f_1.

- The residual service curve can potentially be negative
- The delay bound does not hold

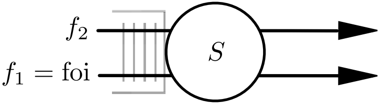
Residuation with a (Min-Plus) Service Curve [BBLC18]

- The residual service curve can potentially be negative
- The delay bound does not hold

Residuation Under (Min-Plus) Service Process



Residuation Under (Min-Plus) Service Process



Theorem

Consider two arrival processes A_1, A_2, that share a server with a min-plus service process S. Further, we assume A_1 to be the arrivals of the foi and an arbitrary multiplexing between the arrivals. Then, the foi sees a residual min-plus service process

$$\xi(s,t) = S(s,t) - A_2(s,t)$$

for all $0 \leq s \leq t$.

Residuation Under (Min-Plus) Service Process

Theorem

Consider two arrival processes A_1, A_2, that share a server with a min-plus service process S. Further, we assume A_1 to be the arrivals of the foi and an arbitrary multiplexing between the arrivals. Then, the foi sees a residual min-plus service process

$$\xi(s,t) = S(s,t) - A_2(s,t)$$

for all $0 \leq s \leq t$.

- The negativity of the service process poses a similar problem

Residuation Under (Min-Plus) Service Process

Theorem

Consider two arrival processes A_1, A_2, that share a server with a min-plus service process S. Further, we assume A_1 to be the arrivals of the foi and an arbitrary multiplexing between the arrivals. Then, the foi sees a residual min-plus service process

$$\xi(s,t) = S(s,t) - A_2(s,t)$$

for all $0 \leq s \leq t$.

- The negativity of the service process poses a similar problem
- In SNC, we still require the server to be strict in multiple flows scenarios

Residuation Under (Min-Plus) Service Process

Residuation Under (Min-Plus) Service Process

- How does one deal with "lazyness" of the service?

Residuation Under (Min-Plus) Service Process

- How does one deal with "lazyness" of the service?
- One needs some sort of guarantee on the incoming traffic

Stochastic Minimum Arrival Guarantees

Generalized Delay Bound

- Guarantee a minimal amount of traffic in the system

Generalized Delay Bound

- Guarantee a minimal amount of traffic in the system
- Achieved through the assumption of a minimal arrival curve $\underline{\alpha}$

Generalized Delay Bound

- Guarantee a minimal amount of traffic in the system
- Achieved through the assumption of a minimal arrival curve $\underline{\alpha}$
- Leverage this information to derive a new delay bound

Generalized Delay Bound

- Guarantee a minimal amount of traffic in the system
- Achieved through the assumption of a minimal arrival curve $\underline{\alpha}$
- Leverage this information to derive a new delay bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

Generalized Delay Bound

- Guarantee a minimal amount of traffic in the system
- Achieved through the assumption of a minimal arrival curve $\underline{\alpha}$
- Leverage this information to derive a new delay bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

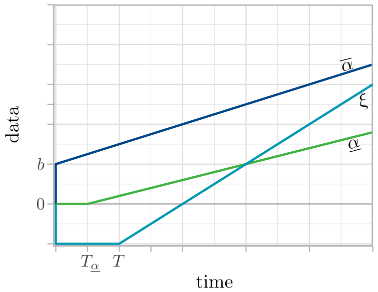
- For proof details, please refer to [HCS24]

Generalized Delay Bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

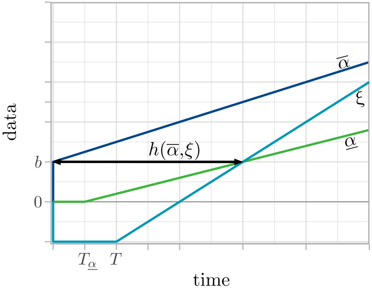


Generalized Delay Bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

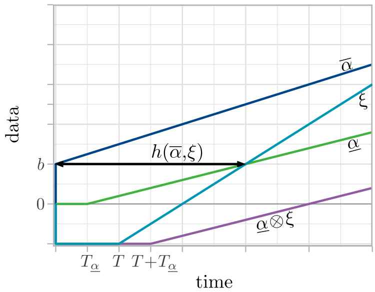


Generalized Delay Bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

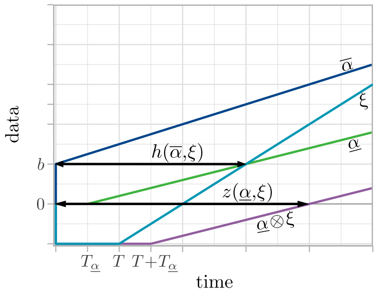


Generalized Delay Bound

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$



Generalized Delay Bound

Minimal Guarantees as the Laplace Transform

- Similar idea as in [HCS24]

Minimal Guarantees as the Laplace Transform

- Similar idea as in [HCS24]
- Guarantee a minimal amount of incoming traffic

Minimal Guarantees as the Laplace Transform

- Similar idea as in [HCS24]
- Guarantee a minimal amount of incoming traffic
- It is represented through the Laplace Transform of the arrival process

Minimal Guarantees as the Laplace Transform

- Similar idea as in [HCS24]
- Guarantee a minimal amount of incoming traffic
- It is represented through the Laplace Transform of the arrival process

Definition

An arrival process $A(s,t)$ is $(\sigma_{\underline{A}}, \rho_{\underline{A}})$-bounded for $\theta > 0$ if for all $0 \leq s \leq t$

$$\phi_{A(s,t)}(-\theta) = \mathrm{E}\left[e^{-\theta A(s,t)}\right] \leq e^{-\theta \rho_{\underline{A}}(-\theta)\cdot(t-s)+\theta\sigma_{\underline{A}}(-\theta)},$$

Minimal Guarantees as the Laplace Transform

Definition

An arrival process $A(s,t)$ is $(\sigma_{\underline{A}}, \rho_{\underline{A}})$-bounded for $\theta > 0$ if for all $0 \leq s \leq t$

$$\phi_{A(s,t)}(-\theta) = \mathrm{E}\left[e^{-\theta A(s,t)}\right] \leq e^{-\theta \rho_{\underline{A}}(-\theta)\cdot(t-s)+\theta\sigma_{\underline{A}}(-\theta)},$$

- In contrast to MGF, we need no extra assumptions

Minimal Guarantees as the Laplace Transform

Definition

An arrival process $A(s,t)$ is $(\sigma_{\underline{A}}, \rho_{\underline{A}})$-bounded for $\theta > 0$ if for all $0 \leq s \leq t$

$$\phi_{A(s,t)}(-\theta) = \mathrm{E}\left[e^{-\theta A(s,t)}\right] \leq e^{-\theta\rho_{\underline{A}}(-\theta)\cdot(t-s)+\theta\sigma_{\underline{A}}(-\theta)},$$

- In contrast to MGF, we need no extra assumptions
- A nontrivial Laplace Transform always exists

Generalized Sample Path Bounds: Backlog

- The sample path backlog bound remains unchanged

Generalized Sample Path Bounds: Backlog

- The sample path backlog bound remains unchanged

Theorem

Consider a flow with arrival process $A(s,t)$ traversing a server with min-plus service process $\xi(s,t)$. The backlog at time $t \geq 0$ is upper bounded by

$$q(t) \leq A \oslash \xi\,(t,t).$$

Generalized Sample Path Bounds: Delay

- The sample path delay bound needs to be generalized

Generalized Sample Path Bounds: Delay

- The sample path delay bound needs to be generalized
- Observe the similarities between both DNC and SNC delay bounds

Generalized Sample Path Bounds: Delay

Theorem

Let an arrival process A traverse a system $\mathcal{S}$. Further, let the arrivals be constrained by maximal arrival curve $\overline{\alpha} \in \mathcal{F}_0^{\uparrow}$ and minimal arrival curve $\underline{\alpha} \in \mathcal{F}_0^{\uparrow}$, and let the system offer a service curve $\xi \in \mathcal{F}_{\leq 0}^{\uparrow}$. The virtual delay $d(t)$ satisfies for all $t \geq 0$

$$d(t) \leq h(\overline{\alpha}, \xi) \vee z(\underline{\alpha}, \xi).$$

Generalized Sample Path Bounds: Delay

Theorem

Consider a flow with arrival process $A(s,t)$ traversing a server with min-plus service process $\xi(s,t)$.
The virtual delay at time $t \geq 0$ is upper bounded by

$$d(t) \leq \inf\left\{s \geq 0 : A \hat{\oslash} \xi(t+s,t) \leq 0 \ AND \ A \hat{\otimes} \xi(t,t+s) \geq 0\right\},$$

$x \hat{\otimes} y(t,t+u) := \inf_{t+1 \leq \tau \leq t+u}\{x(t,\tau) + y(\tau,t+u)\}$ *and*
$x \hat{\oslash} y(t+u,t) := \sup_{0 \leq \tau \leq t}\{x(\tau,t) - y(\tau,t+u)\}$.

Stochastically Generalized Delay Bound

Theorem

Let $\theta > 0$. Suppose we have an arrival process A that is (σ_A, ρ_A)-bounded as well as $(\sigma_{\underline{A}}, \rho_{\underline{A}})$-lower-bounded and a service process ξ that is (σ_ξ, ρ_ξ)-bounded. Additionally, we require the arrivals and the service to be independent. Further, we assume stability condition

$$\rho_A(\theta) < \rho_\xi(-\theta).$$

Let $T \geq 0$. For the virtual delay, it holds for all $t \geq 0$ that

$$\mathrm{P}(d(t) > T) \leq \textit{Standard} + \textit{Penalty}.$$

Stochastically Generalized Delay Bound

Theorem

Let $\theta > 0$. Suppose we have an arrival process A that is (σ_A, ρ_A)-bounded as well as $(\sigma_{\underline{A}}, \rho_{\underline{A}})$-lower-bounded and a service process ξ that is (σ_ξ, ρ_ξ)-bounded. Additionally, we require the arrivals and the service to be independent. Further, we assume stability condition

$$\rho_A(\theta) < \rho_\xi(-\theta).$$

Let $T \geq 0$. For the virtual delay, it holds for all $t \geq 0$ that

$$\begin{aligned}
&\mathrm{P}(d(t) > T) \\
\leq & e^{-\theta\rho_\xi(-\theta)T} e^{\theta\sigma_\xi(-\theta)} \cdot \left(e^{\theta\sigma_A(\theta)} \frac{1}{1 - e^{\theta\left(\rho_A(\theta) - \rho_\xi(-\theta)\right)}} \right. \\
& \left. + e^{\theta\sigma_{\underline{A}}(-\theta)} e^{-\theta\left(\rho_{\underline{A}}(-\theta) - \rho_\xi(-\theta)\right)} \cdot \frac{1 - e^{-\theta\left(\rho_{\underline{A}}(-\theta) - \rho_\xi(-\theta)\right)T}}{1 - e^{-\theta\left(\rho_{\underline{A}}(-\theta) - \rho_\xi(-\theta)\right)}} \right).
\end{aligned}$$

Stochastically Generalized Delay Bound

- How high is the penalty?

Stochastically Generalized Delay Bound

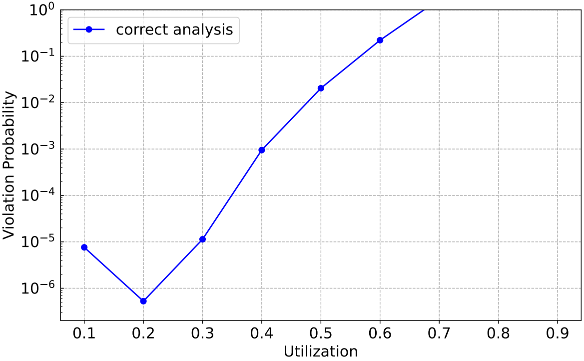
Applications

Toy Example (D/M/1)

- We set the arrival rate $\lambda = 1$ and the delay at 4

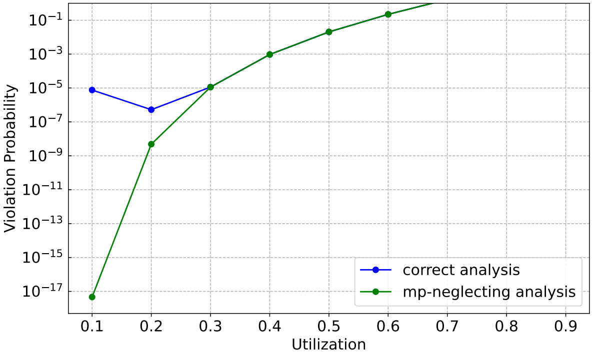
Toy Example (D/M/1)

- We set the arrival rate $\lambda = 1$ and the delay at 4



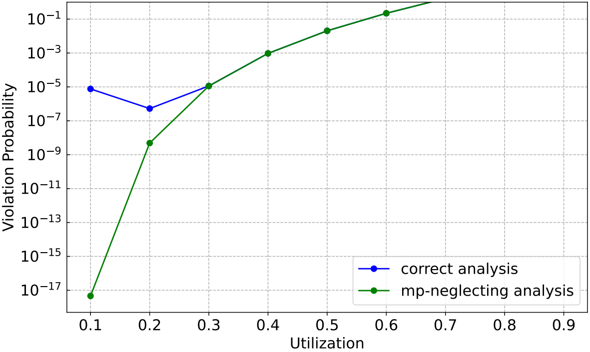
Toy Example (D/M/1)

- We set the arrival rate $\lambda = 1$ and the delay at 4



Toy Example (D/M/1)

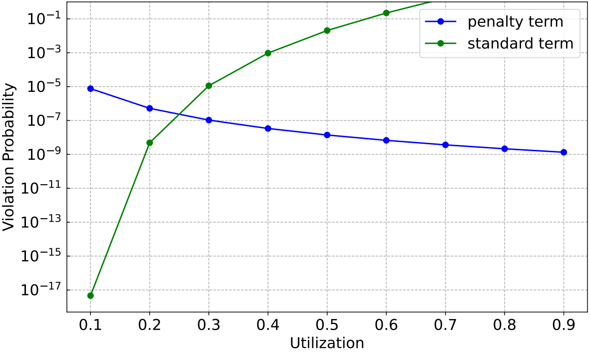
- We set the arrival rate $\lambda = 1$ and the delay at 4



- The bounds are converging from 30% utilization onward

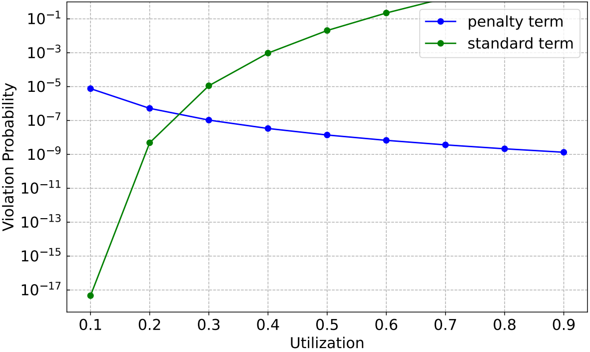
Toy Example (Penalty vs. Standard)

- We set $\lambda = 1$ and the delay at 4



Toy Example (Penalty vs. Standard)

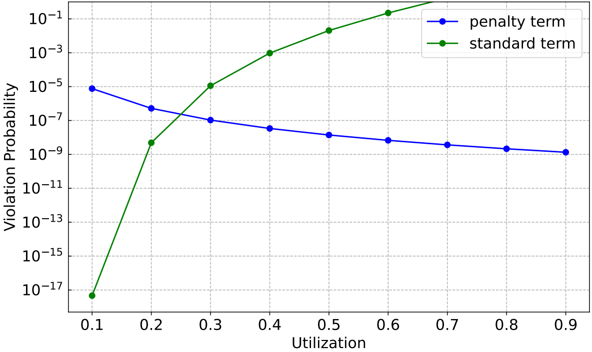
- We set $\lambda = 1$ and the delay at 4



- The standard term rises rapidly

Toy Example (Penalty vs. Standard)

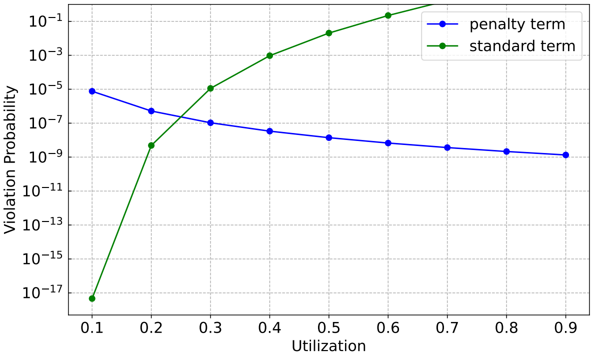
- We set $\lambda = 1$ and the delay at 4



- The standard term rises rapidly
- Our bound decreases with utilization

Toy Example (Penalty vs. Standard)

- We set $\lambda = 1$ and the delay at 4



- The standard term rises rapidly
- Our bound decreases with utilization
- Higher utilization leads to tighter minimum arrival guarantees

Conclusion

Conclusion

- Extended SNC to support a broader class of service processes

Conclusion

- Extended SNC to support a broader class of service processes
- Generalized delay bounds by leveraging minimum arrival guarantees

Conclusion

- Extended SNC to support a broader class of service processes
- Generalized delay bounds by leveraging minimum arrival guarantees
- Released the strictness requirement in multiple flow scenarios

Conclusion

- Extended SNC to support a broader class of service processes
- Generalized delay bounds by leveraging minimum arrival guarantees
- Released the strictness requirement in multiple flow scenarios
- Laplace bounds come for free in this framework

Conclusion

- Extended SNC to support a broader class of service processes
- Generalized delay bounds by leveraging minimum arrival guarantees
- Released the strictness requirement in multiple flow scenarios
- Laplace bounds come for free in this framework
- Additionally, derived new Laplace bounds for Markov-modulated arrival processes

Anne Bouillard, Marc Boyer, and Euriell Le Corronc.
Deterministic Network Calculus: From Theory to Practical Implementation.
John Wiley & Sons, 2018.

Anja Hamscher, Vlad-Cristian Constantin, and Jens B Schmitt.
Extending network calculus to deal with min-plus service curves in multiple flow scenarios.
In *2024 IEEE 30th Real-Time and Embedded Technology and Applications Symposium (RTAS)*, pages 95–107. IEEE, 2024.

Question Time

Question Time

Laplace Bound for Markov-Modulated Arrival Processes

Theorem

Assume a Markov-modulated arrival process A with the finite state space S. The Markov chain is described by its state space S and transition matrix $T = [t_{ij}]$ such that $t_{ij} > 0$ for all $i, j \in S$. We define the increments of the arrival process $a(t) = X_{Y(t)}(t)$, where $X_i(t), i \in S$, is an i.i.d. process with existing MGF, and denote by $E \in \mathrm{Diag}(S)$ the matrix with entries $E_i := E_{ii} := \mathsf{E}\left[e^{-\theta a(t)} \mid Y_t = i\right]$ for all states $i \in S$. Then it holds that the Laplace transform of A is (σ, ρ)-bounded with

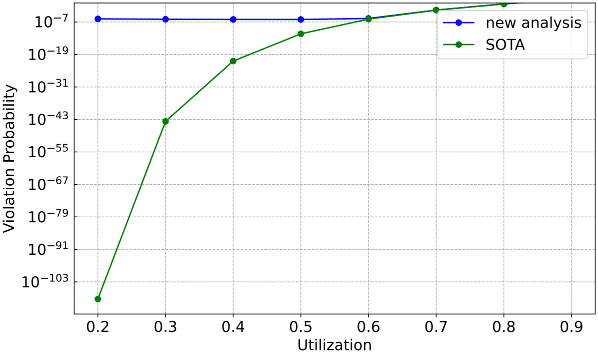
$$\sigma(-\theta) = \frac{1}{\theta} \log\left(\left(\max_{i \in S} E_i\right) \cdot \frac{\max_{k \in S} \bar{x}_k}{\min_{k \in S} \bar{x}_k} \cdot \frac{1}{\mathrm{sp}(ET)}\right),$$

$$\rho(-\theta) = -\frac{1}{\theta} \log\left(\mathrm{sp}(ET)\right),$$

where $\bar{x}$ is a positive eigenvector of ET.

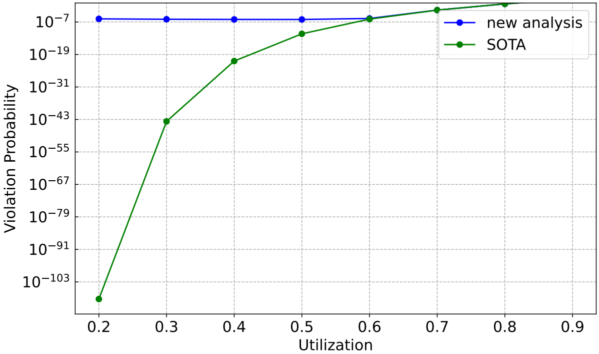
Numerical Evaluation of MM Truncated Normal Arrivals (MMTN)

- MMTN arrival processes offer modelling flexibility
- The parameters of MMTN are set as $\mu = 0.5, \sigma = 1$ and $t_{11} = 0.8, t_{22} = 0.2$
- Target delay is fixed at 10



Numerical Evaluation of MM Truncated Normal Arrivals (MMTN)

- MMTN arrival processes offer modelling flexibility
- The parameters of MMTN are set as $\mu = 0.5, \sigma = 1$ and $t_{11} = 0.8, t_{22} = 0.2$
- Target delay is fixed at 10



- The new bounds increase more gradually