Timescaling results for Markov modulated infinite-server systems and OU processes

Michel Mandjes\textsuperscript{1,2,3}

\textsuperscript{1}Korteweg-de Vries Institute for Mathematics, University of Amsterdam
\textsuperscript{2}CWI, Amsterdam
\textsuperscript{3}Eurandom, Eindhoven

Newton Institute, Cambridge, August 2013

Workshop: ‘Modern probabilistic techniques for design and analysis of stochastic systems and networks’
HOW IT ALL STARTED

Frank Bruggeman (CWI) – biomathematics
HOW IT ALL STARTED

- mRNA is generated according to a Poisson process, parallel decay after generally distributed amount of time.
HOW IT ALL STARTED

- mRNA is generated according to a Poisson process, parallel decay after generally distributed amount of time.
- Complication: arrival rate alternates between ‘high’ and ‘low’.
HOW IT ALL STARTED

- mRNA is generated according to a Poisson process, parallel decay after generally distributed amount of time.
- Complication: arrival rate alternates between ‘high’ and ‘low’.
- Markov-modulated M/G/∞.
HOW IT ALL STARTED

- mRNA is generated according to a Poisson process, parallel decay after generally distributed amount of time.
- Complication: arrival rate alternates between ‘high’ and ‘low’.
- Markov-modulated M/G/∞.
- Little literature.
HOW IT ALL STARTED

- mRNA is generated according to a Poisson process, parallel decay after generally distributed amount of time.
- Complication: arrival rate alternates between ‘high’ and ‘low’.
- Markov-modulated M/G/∞.
- Little literature.
- Our goal: explicit timescale-related limits.
OVERVIEW

- Components of the models considered: infinite-server queues, Ornstein-Uhlenbeck processes, and Markov modulation.
OVERVIEW

- Components of the models considered: infinite-server queues, Ornstein-Uhlenbeck processes, and Markov modulation


- Part II: time-scaling results for Markov-modulated infinite-server queue: large deviations (joint work with Joke Blom, Koen de Turck)

- Part III: time-scaling results for Markov-modulated Ornstein-Uhlenbeck processes: CLT, large deviations (joint work with Gang Huang, Koen de Turck, Peter Spreij)

- Part IV: multiple coupled infinite-server systems (preliminary! – joint work with Koen de Turck, Peter Taylor)
OVERVIEW

- Components of the models considered: infinite-server queues, Ornstein-Uhlenbeck processes, and Markov modulation


- Part II: time-scaling results for Markov-modulated infinite-server queue: large deviations (joint work with Joke Blom, Koen de Turck)

- Part III: time-scaling results for Markov-modulated Ornstein-Uhlenbeck processes: CLT, large deviations (joint work with Gang Huang, Koen de Turck, Peter Spreij)

- Part IV: multiple coupled infinite-server systems (preliminary! – joint work with Koen de Turck, Peter Taylor)
Overview

- Components of the models considered: infinite-server queues, Ornstein-Uhlenbeck processes, and Markov modulation


- Part II: time-scaling results for Markov-modulated infinite-server queue: large deviations (joint work with Joke Blom, Koen de Turck)

- Part III: time-scaling results for Markov-modulated Ornstein-Uhlenbeck processes: CLT, large deviations (joint work with Gang Huang, Koen de Turck, Peter Spreij)

- Part IV: multiple coupled infinite-server systems (preliminary! – joint work with Koen de Turck, Peter Taylor)
OVERVIEW

- Components of the models considered: infinite-server queues, Ornstein-Uhlenbeck processes, and Markov modulation


- Part II: time-scaling results for Markov-modulated infinite-server queue: large deviations (joint work with Joke Blom, Koen de Turck)

- Part III: time-scaling results for Markov-modulated Ornstein-Uhlenbeck processes: CLT, large deviations (joint work with Gang Huang, Koen de Turck, Peter Spreij)

- Part IV: multiple coupled infinite-server systems (preliminary! – joint work with Koen de Turck, Peter Taylor)
INFINITE-SERVER QUEUE

\( M(t) \) lives on \( \{0, 1, 2 \ldots \} \).

- Rate up (from \( i \) to \( i + 1 \)) is \( \lambda \),
- rate down (from \( i \) to \( i - 1 \)) is \( i \mu \).

Fairly complete analysis is possible: steady-state, transient, various performance metrics, etc.
Now impose *Markov modulation*. Let $X(t)$ be (*background process of dimension* $d$).

- $(X(t))_{t \geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$. 

**INFINITE-SERVER QUEUE**
Now impose Markov modulation. Let $X(t)$ be (background process of dimension $d$).

- $(X(t))_{t \geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$.
- Transition rates: $Q = (q_{ij})_{i,j=1}^{d}$, (unique) invariant distribution: $\pi$. 
INFINITE-SERVER QUEUE

Now impose *Markov modulation*. Let $X(t)$ be *(background process of dimension $d$)*.

- $(X(t))_{t \geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$.
- Transition rates: $Q = (q_{ij})_{i,j=1}^{d}$, (unique) invariant distribution: $\pi$.

Let $\lambda$ and $\mu$ be non-negative $d$-dimensional vectors.
INFINITE-SERVER QUEUE

$M(t)$ lives on \{0, 1, 2 \ldots\}.

- Rate up (from $i$ to $i + 1$) is $\lambda_{X(t)}$.
- Rate down (from $i$ to $i - 1$) is $i \mu_{X(t)}$.
INFINITE-SERVER QUEUE

$M(t)$ lives on $\{0, 1, 2 \ldots \}$.

- Rate up (from $i$ to $i + 1$) is $\lambda X(t)$,
- rate down (from $i$ to $i - 1$) is $i \mu X(t)$.

Useful model in e.g. modelling of communication networks (infinite-server model as approximation of many-server model)
INFINITE-SERVER QUEUE

$M(t)$ lives on $\{0, 1, 2 \ldots \}$.

- Rate up (from $i$ to $i + 1$) is $\lambda X(t)$,
- rate down (from $i$ to $i - 1$) is $i \mu X(t)$.

Useful model in e.g. modelling of communication networks (infinite-server model as approximation of many-server model) but also in e.g. biology (generation and decay of mRNA in cells).
INFINITE-SERVER QUEUE

Relatively little number of papers available (\(<\ 30\ldots\)) for resulting model (MMIS – Markov-Modulated Infinite-Server).

Available results for Markov-modulated infinite-server queue typically in terms of \(^d\)-dimensional system of (partial) differential equations to describe \(M(t)\) and stationary counterpart, \(M\).

Recursive scheme to determine all moments; for transient moments in all steps non-homogeneous system of linear differential equations needs to be solved.

See papers by O'Cinneide/Purdue, Keilson/Servi, Adan/Fralix, D'Auria, ...
INFINITE-SERVER QUEUE

Relatively little number of papers available (< 30...) for resulting model (MMIS – Markov-Modulated Infinite-Server).

Available results for Markov-modulated infinite-server queue typically in terms of

- $d$-dimensional system of (partial) differential equations to describe pgf of $M(t)$ and stationary counterpart, $M$.
- Recursive scheme to determine all moments; for transient moments in all steps non-homogeneous system of linear differential equations needs to be solved.
INFINITE-SERVER QUEUE

Relatively little number of papers available (< 30...) for resulting model (MMIS – Markov-Modulated Infinite-Server).

Available results for Markov-modulated infinite-server queue typically in terms of

- $d$-dimensional system of (partial) differential equations to describe pgf of $M(t)$ and stationary counterpart, $M$.
- Recursive scheme to determine all moments; for transient moments in all steps non-homogeneous system of linear differential equations needs to be solved.

See papers by O’Cinneide/Purdue, Keilson/Servi, Adan/Fralix, D’Auria, ...
INFINITE-SERVER QUEUE

For Markov-modulated single-server queue a lot is known (Neuts): stationary number in the system follows a matrix-geometric distribution (generalization of M/M/1).
INFINITE-SERVER QUEUE

For Markov-modulated single-server queue a lot is known (Neuts): stationary number in the system follows a matrix-geometric distribution (generalization of M/M/1).

Therefore in the context of MMIS one would naïvely expect a matrix-Poisson distribution (generalization of M/M/∞)…
INFINITE-SERVER QUEUE

For Markov-modulated single-server queue a lot is known (Neuts): stationary number in the system follows a matrix-geometric distribution (generalization of M/M/1).

Therefore in the context of MMIS one would na"ively expect a matrix-Poisson distribution (generalization of M/M/∞)...

but this is not true.
INFINITE-SERVER QUEUE

In above model (referred to as Model I) the transition rates depend on the current state of the background process.

\[ M(t) = \int_0^t \lambda X(s) e^{-\mu X(s)(t-s)} ds \]
In above model (referred to as Model I) the transition rates depend on the current state of the background process.

Alternative model (Model II): service times are sampled upon arrival. D’Auria: if $M(0) = 0$, then $M(t)$ has a Poisson distribution with random parameter

$$\int_0^t \lambda \chi(s) e^{-\mu \chi(s)(t-s)} ds.$$
ORNSTEIN AND UHLENBECK
OU PROCESS

- Stochastic differential equation

\[ dM(t) = (\alpha - \gamma M(t))dt + \sigma dB(t), \]

where \( \alpha, \gamma, \sigma > 0 \), \( B(t) \) is a standard Brownian motion.
OU PROCESS

- Stochastic differential equation

\[ dM(t) = (\alpha - \gamma M(t))dt + \sigma dB(t), \]

where \( \alpha, \gamma, \sigma > 0 \), \( B(t) \) is a standard Brownian motion.

- Similarity with the infinite-server queue. There jobs are generated according to a Poisson process of rate \( \lambda \). They remain in system \( \exp(\mu) \) time; they don’t “see” each other, so departure rate is \( \mu \) multiplied by number of jobs present.
OU PROCESS

It is elementary (bit of stochastic calculus) to obtain that $M(t)$ has a Normal distribution with
OU PROCESS

It is elementary (bit of stochastic calculus) to obtain that \( M(t) \) has a Normal distribution with

- mean:
  \[
  \mathbb{E}M(t) = \mathbb{E}M(0)e^{-\gamma t} + \frac{\alpha}{\gamma}(1 - e^{-\gamma t}),
  \]

OU is a Markovian, Gaussian process, that is mean-reverting (towards the limiting mean \( \frac{\alpha}{\gamma} \)).
OU PROCESS

It is elementary (bit of stochastic calculus) to obtain that $M(t)$ has a Normal distribution with

- mean:
  \[ \mathbb{E} M(t) = \mathbb{E} M(0)e^{-\gamma t} + \frac{\alpha}{\gamma} (1 - e^{-\gamma t}), \]

- variance:
  \[ \text{Var } M(t) = \text{Var} \left( \sigma \int_0^t e^{-\gamma(t-s)} dB(s) \right) = \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma t}). \]
OU PROCESS

It is elementary (bit of stochastic calculus) to obtain that $M(t)$ has a Normal distribution with

- mean:
  $$\mathbb{E} M(t) = \mathbb{E} M(0) e^{-\gamma t} + \frac{\alpha}{\gamma} (1 - e^{-\gamma t}),$$

- variance:
  $$\text{Var} M(t) = \text{Var} \left( \sigma \int_0^t e^{-\gamma (t-s)} dB(s) \right) = \frac{\sigma^2}{2\gamma} (1 - e^{-2\gamma t}).$$

OU is a Markovian, Gaussian process, that is mean-reverting (towards the limiting mean $\alpha/\gamma$).
OU PROCESS

Again: consider effect of Markov modulation: parameters $\alpha, \gamma, \sigma > 0$ have values $\alpha_i, \gamma_i, \sigma_i > 0$ when independent background Markov chain is in state $i$. 
MODEL: MMOU

- $(X(t))_{t\geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$. 

Transition rates: $Q = (q_{ij})_{i,j=1}^d$, (unique) invariant distribution: $\pi$.

Now we suppose that the process $X(t)$ modulates an Ornstein-Uhlenbeck process: while $X(t)$ in state $i$, the process $(M(t))_{t\geq 0}$ behaves as an Ornstein-Uhlenbeck process $U_i(t)$ with parameters $\alpha_i, \gamma_i, \sigma_i$, independently of the 'background process' $X(t)$.

Hence, $M(t)$ obeys the following SDE:
$$dM(t) = (\alpha_{X(t)} - \gamma_{X(t)} M(t)) dt + \sigma_{X(t)} dB(t);$$
where $(B(t))_{t\geq 0}$ standard BM independent of $(X(t))_{t\geq 0}$.

Queueing: Markov modulation — Finance: regime switching.
MODEL: MMOU

- $(X(t))_{t \geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$.
- Transition rates: $Q = (q_{ij})_{i,j=1}^d$, (unique) invariant distribution: $\pi$.

Now we suppose that the process $X(\cdot)$ modulates an Ornstein-Uhlenbeck process: while $X(\cdot)$ in state $i$, the process $(M(t))_{t \geq 0}$ behaves as an Ornstein-Uhlenbeck process $U_i(\cdot)$ with parameters $\alpha_i, \gamma_i, \sigma_i$, independently of the 'background process' $X(\cdot)$.

Hence, $M(\cdot)$ obeys the following SDE:

$$dM(t) = (\alpha X(t) - \gamma X(t) M(t)) \, dt + \sigma X(t) \, dB(t); \quad (B(t))_{t \geq 0} \text{ standard BM independent of } (X(t))_{t \geq 0}.$$
MODEL: MMOU

- \( (X(t))_{t \geq 0} \): irreducible, Markov process on \( \{1, \ldots, d\} \).
- Transition rates: \( Q = (q_{ij})_{i,j=1}^{d} \), (unique) invariant distribution: \( \pi \).
- Now we suppose that the process \( X(\cdot) \) modulates an Ornstein-Uhlenbeck process: while \( X(\cdot) \) in state \( i \), the process \( (M(t))_{t \geq 0} \) behaves as an Ornstein-Uhlenbeck process \( U_i(\cdot) \) with parameters \( \alpha_i, \gamma_i, \) and \( \sigma_i \), independently of the ‘background process’ \( X(\cdot) \).
MODEL: MMOU

- $(X(t))_{t\geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$.
- Transition rates: $Q = (q_{ij})_{i,j=1}^{d}$, (unique) invariant distribution: $\pi$.

Now we suppose that the process $X(\cdot)$ modulates an Ornstein-Uhlenbeck process: while $X(\cdot)$ in state $i$, the process $(M(t))_{t\geq 0}$ behaves as an Ornstein-Uhlenbeck process $U_i(\cdot)$ with parameters $\alpha_i$, $\gamma_i$, and $\sigma_i$, independently of the ‘background process’ $X(\cdot)$.

Hence, $M(\cdot)$ obeys the following SDE:

$$dM(t) = (\alpha X(t) - \gamma X(t) M(t))dt + \sigma X(t) dB(t);$$

$(B(t))_{t\geq 0}$ standard BM independent of $(X(t))_{t\geq 0}$. 
MODEL: MMOU

- $(X(t))_{t \geq 0}$: irreducible, Markov process on $\{1, \ldots, d\}$.
- Transition rates: $Q = (q_{ij})_{i,j=1}^d$, (unique) invariant distribution: $\pi$.
- Now we suppose that the process $X(\cdot)$ modulates an Ornstein-Uhlenbeck process: while $X(\cdot)$ in state $i$, the process $(M(t))_{t \geq 0}$ behaves as an Ornstein-Uhlenbeck process $U_i(\cdot)$ with parameters $\alpha_i, \gamma_i, \text{ and } \sigma_i$, independently of the ‘background process’ $X(\cdot)$.
- Hence, $M(\cdot)$ obeys the following SDE:

$$dM(t) = (\alpha X(t) - \gamma X(t) M(t))dt + \sigma X(t) dB(t);$$

$(B(t))_{t \geq 0}$ standard BM independent of $(X(t))_{t \geq 0}$.

- Queueing: Markov modulation — Finance: regime switching.
Denote by $X$ the path $(X(s), s \in [0, t])$. $(M(t) \mid X)$ has a Normal distribution with random parameters $\mathbb{M}$ and $\mathbb{S}$ given by

$$M := \mathbb{E}(M(t) \mid X) = m_0 \exp \left( - \int_0^t \gamma X(s) \, ds \right) + \int_0^t \exp \left( - \int_s^t \gamma X(r) \, dr \right) \alpha X(s) \, ds$$

and

$$S := \mathbb{V} \text{ar}(M(t) \mid X) = \int_0^t \exp \left( -2 \int_s^t \gamma X(r) \, dr \right) \sigma_X^2 X(s) \, ds.$$
Mou: conditional mean and variance

Denote by $X$ the path $(X(s), s \in [0, t])$. $(M(t) \mid X)$ has a Normal distribution with random parameters $M$ and $S$ given by

$$M := \mathbb{E}(M(t) \mid X)$$

$$= m_0 \exp \left( - \int_0^t \gamma X(s) ds \right) + \int_0^t \exp \left( - \int_s^t \gamma X(r) dr \right) \alpha X(s) ds$$

and

$$S := \text{Var}(M(t) \mid X) = \int_0^t \exp \left( -2 \int_s^t \gamma X(r) dr \right) \sigma^2 X(s) ds.$$

Similarity with corresponding result for MMIS queue by D’Auria: there number of jobs in system has a Poisson distribution with random parameter.
Part I
MARKOV MODULATED INFINITE-SERVER (MMIS)
Central Limit Theorems
MMIS: unconditional mean and variance

For ease (in this presentation): take $\mu_i$ identical so that Model I and Model II coincide.
MMIS: unconditional mean and variance

For ease (in this presentation): take $\mu_i$ identical so that Model I and Model II coincide. (Other models technically more demanding.)
**MMIS: unconditional mean and variance**

For ease (in this presentation): take $\mu_i$ identical so that Model I and Model II coincide. (Other models technically more demanding.)

Straightforward (for instance from Poisson-with-random-mean representation):

$$\mathbb{E}M(t) = \sum_{j=1}^{d} \pi_i \frac{\lambda_i}{\mu} (1 - e^{-\mu t}),$$

if background process starts in stationarity.
MMIS: unconditional mean and variance

Variance can be computed with law of total variance:

\[ \Var M(t) = \mathbb{E}(\Var(M(t) \mid X)) + \Var(\mathbb{E}(M(t) \mid X)), \]

with \( X \equiv (X(s))_{s \in [0,t]} \).
MMIS: unconditional mean and variance

Variance can be computed with law of total variance:

$$\text{Var} M(t) = \mathbb{E}(\text{Var}(M(t) \mid X)) + \text{Var}(\mathbb{E}(M(t) \mid X)),$$

with \( X \equiv (X(s))_{s \in [0,t]} \).

Clearly,

$$\mathbb{E}(\text{Var}(M(t) \mid X)) = \mathbb{E} M(t) = \sum_{j=1}^{d} \pi_i \frac{\lambda_i}{\mu} (1 - e^{-\mu t}).$$
MMIS: unconditioned mean and variance

\[
\text{Var}(\mathbb{E}(M(t) \mid X)) = \text{Var} \left( \int_0^t \lambda_X(s) e^{-\mu(t-s)} \, ds \right)
\]

\[
= \int_0^t \int_0^t \text{Cov} \left( \lambda_X(s) e^{-\mu(t-s)}, \lambda_X(u) e^{-\mu(t-u)} \right) \, ds \, du
\]

\[
= \sum_{i,j=1}^d \lambda_i \lambda_j \int_0^t \int_0^t e^{-\mu(t-s)} e^{-\mu(t-u)} \text{Cov} \left( 1_{\{X(s)=i\}}, 1_{\{X(u)=j\}} \right) \, ds \, du.
\]
MMIS: unconditional mean and variance

\[
\text{Var}(\mathbb{E}(M(t) \mid X)) = \text{Var} \left( \int_0^t \lambda_X(s) e^{-\mu(t-s)} ds \right) \\
= \int_0^t \int_0^t \text{Cov} \left( \lambda_X(s) e^{-\mu(t-s)}, \lambda_X(u) e^{-\mu(t-u)} \right) ds du \\
= \sum_{i,j=1}^d \lambda_i \lambda_j \int_0^t \int_0^t e^{-\mu(t-s)} e^{-\mu(t-u)} \text{Cov} \left( 1 \{X(s)=i\}, 1 \{X(u)=j\} \right) ds du.
\]

Starting in stationarity:

\[
\sum_{i,j=1}^d \lambda_i \lambda_j \int_0^t \int_0^u e^{-\mu(t-s)} e^{-\mu(t-u)} \pi_i(p_{ij}(u - s) - \pi_j) ds du \\
\quad + \sum_{i,j=1}^d \lambda_i \lambda_j \int_0^t \int_u^t e^{-\mu(t-s)} e^{-\mu(t-u)} \pi_i(p_{ij}(u - s) - \pi_j) ds du.
\]
MMIS: *unconditional* mean and variance

Deviation matrix:

\[ D_{ij} := \int_{0}^{\infty} (p_{ij}(t) - \pi_j) \, dt. \]
MMIS: unconditional mean and variance

Deviation matrix:

\[ D_{ij} := \int_{0}^{\infty} (p_{ij}(t) - \pi_j) dt. \]

Perform timescaling \( \lambda \mapsto \lambda N \), and \( Q \mapsto QN^f \).
MMIS: unconditional mean and variance

Deviation matrix:

\[ D_{ij} := \int_{0}^{\infty} (p_{ij}(t) - \pi_j) \, dt. \]

Perform timescaling $\lambda \mapsto \lambda N$, and $Q \mapsto QN^f$.

Elementary calculations for stationary number in system:

\[
\var M^{(N)} \sim N \sum_{j=1}^{d} \pi_i \frac{\lambda_i}{\mu} + N^{2-f} \sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.
\]
MMIS: *un*conditional mean and variance

Interesting dichotomy:

\[ V_{\text{ar}}(N) \sim N(\varrho), \quad \varrho := d \sum_{j=1}^{\pi} \lambda^i \lambda^j \mu^D_{ij}. \]

'Local equilibria' (?) Can this be phrased as CLT? – apparently the right scaling is \( N^{\gamma}, \) with \( \gamma := \max\{\frac{1}{2}, 1 - f^2\}. \)
MMIS: unconditional mean and variance

Interesting dichotomy:

- If $f > 1$ the variance essentially equals

$$\text{Var} M^{(N)} \sim N \phi, \quad \text{where} \quad \phi := \sum_{j=1}^{d} \pi_i \lambda_i / \mu.$$

The system behaves ‘Poissonian’: background process moves faster than arrival process.
MMIS: unconditional mean and variance

Interesting dichotomy:

- If \( f > 1 \) the variance essentially equals

\[
\text{Var} M^{(N)} \sim N \varrho, \quad \text{where} \quad \varrho := \sum_{j=1}^{d} \pi_i \frac{\lambda_j}{\mu}.
\]

The system behaves ‘Poissonian’: background process moves faster than arrival process.

- If \( f < 1 \) the variance essentially equals

\[
\text{Var} M^{(N)} \sim N^{2-f} \sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.
\]

‘Local equilibria’ (?)
MMIS: unconditional mean and variance

Interesting dichotomy:

- If $f > 1$ the variance essentially equals

$$\text{Var} M^{(N)} \sim N \varrho, \quad \text{where} \quad \varrho := \sum_{j=1}^{d} \pi_i \frac{\lambda_i}{\mu}.$$ 

The system behaves ‘Poissonian’: background process moves faster than arrival process.

- If $f < 1$ the variance essentially equals

$$\text{Var} M^{(N)} \sim N^{2-f} \sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.$$ 

‘Local equilibria’ (?)

Can this be phrased as CLT? – apparently the right scaling is $N^\gamma$, with $\gamma := \max\{\frac{1}{2}, 1 - \frac{f}{2}\}$.
**MMIS: CLT**

Procedure (for steady-state, same can be done for transient):

- Set up a DE for the PGF of $M(N)$.
- Transform this into a DE for the MGF of $M(N) - N\rho N\gamma$.
- Manipulate this expression and let $N \to \infty$.
- Observe that we obtain a Gaussian limit.
**MMIS: CLT**

Procedure (for steady-state, same can be done for transient):
- Set up a DE for the PGF of $M^{(N)}$. 
- Transform this into a DE for the MGF of $M^{(N)} - N\rho N\gamma$. 
- Manipulate this expression and let $N \to \infty$. 
- Observe that we obtain a Gaussian limit.
MMIS: CLT

Procedure (for steady-state, same can be done for transient):

- Set up a DE for the PGF of $M^{(N)}$.
- Transform this into a DE for the MGF of $M^{(N)} - N \varrho N^\gamma$.
- Manipulate this expression and let $N \to \infty$.
- Observe that we obtain a Gaussian limit.

\[
\frac{M^{(N)} - N \varrho}{N^\gamma}.
\]
MMIS: CLT

Procedure (for steady-state, same can be done for transient):

- Set up a DE for the PGF of $M^{(N)}$.
- Transform this into a DE for the MGF of
  \[
  \frac{M^{(N)} - N^Q}{N^\gamma}.
  \]
- Manipulate this expression and let $N \to \infty$. 
MMIS: CLT

Procedure (for steady-state, same can be done for transient):

- Set up a DE for the PGF of $M^{(N)}$.
- Transform this into a DE for the MGF of $M^{(N)} - N_{\varrho}N^{\gamma}$.
- Manipulate this expression and let $N \to \infty$.
- Observe that we obtain a Gaussian limit.
MMIS: CLT

First characterize invariant distribution \((p_k^{(N)})_{k=0}^{\infty}\), where \(p_k\) is \(d\)-dimensional row-vector, defined by

\[
[p_k^{(N)}]_j := \mathbb{P}(M^{(N)} = k, X^{(N)} = j).
\]

The (row-vector-)pgf \(p^{(N)}(z)\) is then given by

\[
p^{(N)}(z) := \sum_{k=0}^{\infty} p_k^{(N)} z^k.
\]
MMIS: CLT

First characterize invariant distribution \((p_k^{(N)})_{k=0}^{\infty}\), where \(p_k\) is \(d\)-dimensional row-vector, defined by

\[
[p_k^{(N)}]_j := \mathbb{P}(M^{(N)} = k, X^{(N)} = j).
\]

The (row-vector-\)pgf \(p^{(N)}(z)\) is then given by

\[
p^{(N)}(z) := \sum_{k=0}^{\infty} p_k^{(N)} z^k.
\]

Elementary (from Kolmogorov equations)

\[
p^{(N)}(z)Q = \frac{(z - 1)}{Nf} \left( (p^{(N)})'(z) \text{diag}\{\mu\} - Np^{(N)}(z) \text{diag}\{\lambda\} \right).
\]
MMIS: CLT

Define:

\[ \Pi := \frac{1}{\pi T} \]

\[ F := D + \Pi \text{ (fundamental matrix).} \]

Standard properties:

\[ QF = FQ = \Pi - I, \]

\[ F_1 = 1, \]

and

\[ \Pi D = D \Pi = 0. \]

Why could one expect deviation matrix showing up here? Let

\[ Z(N)(t) := \frac{nf}{2} (\int_0^t \{ X(N)(s) = i \} ds - \pi it). \]

Then

\[ Z(N)(t) \]

converges to a zero mean Gaussian distribution with covariance matrix

\[ Ct, \]

with

\[ C := D^T \text{diag} \{ \pi \} + \text{diag} \{ \pi \} D. \]

(Cf. e.g. Thm. 4.11 book Asmussen; also Kurtz/Protter, ...).
MMIS: CLT

Define:

- $\Pi := \mathbf{1}\pi^T$. 

Why could one expect deviation matrix showing up here? Let

$$Z(N(i)(t)) := \frac{\mathbf{1}}{2} \left( \int_0^t \{X(N(s)) = i\} ds - \pi it \right).$$

Then $Z(N)(t)$ converges to a zero mean Gaussian distribution with covariance matrix $C_t$, where

$$C := \mathbf{D}^T \text{diag}\{\pi\} + \text{diag}\{\pi\} \mathbf{D}.$$  

(Cf. e.g. Thm. 4.11 book Asmussen; also Kurtz/Protter, ...).
Define:
- $\Pi := 1\pi^T$.
- $F := D + \Pi$ (fundamental matrix).
MMIS: CLT

Define:

- \( \Pi := 1 \pi^T \).
- \( F := D + \Pi \) (fundamental matrix).
- Standard properties: \( QF = FQ = \Pi - I \), \( F1 = 1 \), and \( \Pi D = D \Pi = 0 \).

Why could one expect deviation matrix showing up here? Let \( Z_N(t) := N \frac{\sqrt{\int_0^t 1 \{ X_N(s) = i\} \, ds}}{t} - \pi i \).

Then \( Z_N(t) \) converges to a zero mean Gaussian distribution with covariance matrix \( C_t \), with \( C_t := D^T \text{diag}\{\pi\} + \text{diag}\{\pi\} D \).

(Cf. e.g. Thm. 4.11 book Asmussen; also Kurtz/Protter, ...).
Define:

- $\Pi := 1\pi^T$.
- $F := D + \Pi$ (fundamental matrix).
- Standard properties: $QF = FQ = \Pi - I, F1 = 1$, and $\Pi D = D\Pi = 0$.

Why could one expect deviation matrix showing up here? Let

$$Z_i^{(N)}(t) := N^{f/2} \left( \int_0^t 1_{\{X^{(N)}(s) = i\}} \, ds - \pi_i t \right).$$

Then $Z^{(N)}(t)$ converges to a zero mean Gaussian distribution with covariance matrix $Ct$, with $C := D^T \text{diag}\{\pi\} + \text{diag}\{\pi\} D$. (Cf. e.g. Thm. 4.11 book Asmussen; also Kurtz/Protter, . . .).
MMIS: CLT

Postmultiplying DE with $F$ yields

$$p^{(N)}(z) = p^{(N)}(z)\Pi$$

$$+ N^{-f}(z - 1) \left[ Np^{(N)}(z)\text{diag}\{\lambda\} - (p^{(N)})'(z)\text{diag}\{\mu\} \right] F.$$
MMIS: CLT

Convert this into DE for MGF $\tilde{p}^{(N)}(\vartheta)$ of centered/normalized version of $M^{(N)}$:

$$
\tilde{p}^{(N)}(\vartheta) = \tilde{p}^{(N)}(\vartheta)\Pi + N^{1-f} \left( z^{(N)}(\vartheta) - 1 \right) \tilde{p}^{(N)}(\vartheta)\text{diag}\{\lambda\} F \\
- N^{1-f} \left( 1 - \frac{1}{z^{(N)}(\vartheta)} \right) \circ \tilde{p}^{(N)}(\vartheta)\text{diag}\{\mu\} F \\
- N^{1-f-\beta/2} \left( 1 - \frac{1}{z^{(N)}(\vartheta)} \right) (\tilde{p}^{(N)})'(\vartheta)\text{diag}\{\mu\} F.
$$

Here: $\beta := \min\{f, 1\}$, and $z := z^{(N)}(\vartheta) := \exp(\vartheta N^{-1+\beta/2})$. 
MMIS: CLT

‘Taylor’ the $z$, and iterate the equation to get rid of all terms that are $o(N^{-f})$:
MMIS: CLT

‘Taylor’ the $z$, and iterate the equation to get rid of all terms that are $o(N^{-f})$:

\[
\tilde{p}^{(N)}(\vartheta) = \tilde{p}^{(N)}(\vartheta)\Pi + \vartheta N^{\beta/2-f} \tilde{p}^{(N)}(\vartheta)\Pi(\text{diag}\{\lambda\} - \varrho\text{diag}\{\mu\})F \\
+ \vartheta^2 N^{\beta-2f} \tilde{p}^{(N)}(\vartheta)\Pi(\text{diag}\{\lambda\} - \varrho\text{diag}\{\mu\})F(\text{diag}\{\lambda\} - \varrho\text{diag}\{\mu\})F \\
+ \frac{\vartheta^2 N^{\beta-1-f}}{2} \tilde{p}^{(N)}(\vartheta)\Pi(\text{diag}\{\lambda\} + \varrho\text{diag}\{\mu\})F \\
- \vartheta N^{-f} (\tilde{p}^{(N)})'(\vartheta)\Pi\text{diag}\{\mu\}F + o(N^{-f})
\]
Goal: transform the coupled system of ODEs in $\tilde{p}^{(N)}(\vartheta)$ into a single-dimensional ODE in terms of $\phi^{(N)}(\vartheta) := \tilde{p}^{(N)}(\vartheta)^T$. 
**MMIS: CLT**

Goal: transform the coupled system of ODEs in $\tilde{p}^{(N)}(\vartheta)$ into a single-dimensional ODE in terms of $\phi^{(N)}(\vartheta) := \tilde{p}^{(N)}(\vartheta) \mathbf{1}$.

Postmultiply by $\mathbf{1} N^f / \vartheta$; realize that $\Pi \mathbf{1} = \mathbf{1}$ and $F \mathbf{1} = \mathbf{1}$. 
**MMIS: CLT**

Goal: transform the coupled system of ODEs in $\tilde{p}^N(\vartheta)$ into a single-dimensional ODE in terms of $\phi^N(\vartheta) := \tilde{p}^N(\vartheta)1$.

Postmultiply by $1N^f/\vartheta$; realize that $\Pi1 = 1$ and $F1 = 1$.

Observe that, from the definition of $\varrho$,

$$
\begin{align*}
\tilde{p}^N(\vartheta)\Pi(\text{diag}\{\lambda\} - \varrho\text{diag}\{\mu\})F1 \\
= \phi^N(\vartheta)\pi^T(\text{diag}\{\lambda\} - \varrho\text{diag}\{\mu\})1 = 0.
\end{align*}
$$
We thus obtain

\[
(\phi^{(N)})'(\vartheta) = \vartheta N^{\beta - f} \phi^{(N)}(\vartheta) \frac{\pi^T (\text{diag}\{\lambda\} - \rho \text{diag}\{\mu\}) F (\text{diag}\{\lambda\} - \rho \text{diag}\{\mu\}) \mathbf{1}}{\mu_\infty} \\
+ \vartheta N^{\beta - 1} \rho \phi^{(N)}(\vartheta) + o(1),
\]

using

\[
(\tilde{\rho}^{(N)})'(\vartheta) \Pi \text{diag}\{\mu\} F \mathbf{1} = \mu (\phi^{(N)})'(\vartheta)
\]

and

\[
\pi^T (\text{diag}\{\lambda\} + \rho \text{diag}\{\mu\}) \mathbf{1} = 2\lambda_\infty.
\]
MMIS: CLT

Remember: $\beta = \min\{f, 1\}$.

Conclude:

▶ If $f < 1$, then only first term RHS matters. Obtain Normal distribution with variance $\sum_{i,j=1}^{\pi} \lambda_i \lambda_j \mu_{ij}$.

▶ If $f > 1$, then only second term RHS matters. Obtain Normal distribution with variance $\sum_{i=1}^{\pi} \lambda_i \mu$.

▶ If $f = 1$ both terms matter.
MMIS: CLT

Remember: $\beta = \min\{f, 1\}$.

Conclude:

- If $f < 1$, then only first term RHS matters. Obtain Normal distribution with variance
  
  \[
  \sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.
  \]

- If $f > 1$, then only second term RHS matters. Obtain Normal distribution with variance

- If $f = 1$ both terms matter.
MMIS: CLT

Remember: $\beta = \min\{f, 1\}$.

Conclude:

- If $f < 1$, then only first term RHS matters. Obtain Normal distribution with variance

$$\sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.$$ 

- If $f > 1$, then only second term RHS matters. Obtain Normal distribution with variance

$$\sum_{i=1}^{d} \pi_i \frac{\lambda_i}{\mu}.$$
MMIS: CLT

Remember: $\beta = \min\{f, 1\}$.

Conclude:

- If $f < 1$, then only first term RHS matters. Obtain Normal distribution with variance

  $$\sum_{i,j=1}^{d} \pi_i \frac{\lambda_i \lambda_j}{\mu} D_{ij}.$$ 

- If $f > 1$, then only second term RHS matters. Obtain Normal distribution with variance

  $$\sum_{i=1}^{d} \pi_i \frac{\lambda_i}{\mu}.$$ 

- If $f = 1$ both terms matter.
Variants of this result for Model I and Model II (work with Blom and De Turck, in progress), and for transient as well as steady-state; previous work with Blom, Kella, and Thorsdottir (QUESTA, 2013) just covered $f > 1$ and Model I.
MMIS: CLT

- Variants of this result for Model I and Model II (work with Blom and De Turck, in progress), and for transient as well as steady-state; previous work with Blom, Kella, and Thorsdottir (QUESTA, 2013) just covered $f > 1$ and Model I.

- Possible to extend this to functional versions (convergence to appropriate OU process). Can be done by writing $M^{(N)}$ as difference of two Poisson processes with random time-change:

$$M^{(N)}(t) = Y_1 \left( \int_0^t N \sum_{i=1}^d \lambda_i I_i^{(N)}(s) ds \right) - Y_2 \left( \int_0^t \mu M^{(N)}(s) ds \right),$$

with $I_i^{(N)}(t)$ the indicator function of $\{X^{(N)}(t) = i\}$ and $Y_1$ and $Y_2$ independent unit rate Poisson processes. Then straightforward application of martingale-CLT (work with Anderson, Blom, De Turck, Thorsdottir, in progress).
Part II
MARKOV MODULATED INFINITE-SERVER QUEUES (MMIS)
Large deviations
MMIS: LD

Under the same scaling, large deviations can be examined. Objective:

\[
\lim_{N \to \infty} \frac{1}{N} \log \mathbb{P} \left( \frac{M^{(N)}(t)}{N} \geq a \right).
\]
Under the same scaling, large deviations can be examined. Objective:

\[ \lim_{N \to \infty} \frac{1}{N} \log \mathbb{P} \left( \frac{M^{(N)}(t)}{N} \geq a \right) \]

- **Stationary case:**
  If \( f > 1 \) rate function looks like that of Poisson random variable with parameter

  \[ \frac{\pi^T \lambda}{\pi^T \mu} \text{ and } \pi^T \varrho \]

  for Model I and II, respectively; here \( \varrho_i := \lambda_i / \mu_i \).
MMIS: LD

Under the same scaling, large deviations can be examined.

Objective:

\[
\lim_{N \to \infty} \frac{1}{N} \log \mathbb{P} \left( \frac{M^{(N)}(t)}{N} \geq a \right).
\]

- Stationary case:
  If \( f > 1 \) rate function looks like that of Poisson random variable with parameter
  \[
  \frac{\pi^T \lambda}{\pi^T \mu}, \quad \text{and} \quad \pi^T \varrho
  \]
  for Model I and II, respectively; here \( \varrho_i := \lambda_i/\mu_i \).

- Similar result for transient case and \( f > 1 \). (See paper with Blom and De Turck, Stoch. Mod.)
Crucially different behavior for $f < 1$ – take for ease $f = 0$ (that is, background process is unscaled) and Model II. Recall: $M^{(N)}(t)$ has a Poisson distribution with parameter

$$N \int_0^t \lambda \chi(s) e^{-\mu \chi(s) (t-s)} ds.$$
Crucially different behavior for $f < 1$ – take for ease $f = 0$ (that is, background process is unscaled) and Model II. Recall: $M^{(N)}(t)$ has a Poisson distribution with parameter

$$N \int_0^t \lambda X(s) e^{-\mu X(s)} (t-s) ds.$$

Intuition: a single path of $X(s)$ determines large deviations.
Intuition: a single path of $X(s)$ determines large deviations.
MMIS: LD

Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$. 
MMIS: LD

Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\frac{\lambda_i}{\mu_i}$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg \max_{f(s)} \left\{ \frac{\lambda_f(s)}{\mu_f(s)} (1 - e^{-\mu_f(s)(t-s)}) \right\}.$$
Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg\max_{f(s)} \left\{ \frac{\lambda(f(s))}{\mu(f(s))} (1 - e^{-\mu(f(s))(t-s)}) \right\}.$$ 

On optimal path, background process may jump at most $d - 1$ times. Idea: maximize parameter of Poisson distribution. (See paper with Blom, OR Letters. Extension to general service times.)
Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg \max_{f(s)} \left\{ \frac{\lambda f(s)}{\mu f(s)} (1 - e^{-\mu f(s) (t-s)}) \right\}.$$ 

On optimal path, background process may jump at most $d - 1$ times. Idea: maximize parameter of Poisson distribution. (See paper with Blom, OR Letters. Extension to general service times.)

Interesting shape of decay rate of $\mathbb{P}(\left|M^{(N)}(t)/N - \rho_t\right| \geq \varepsilon)$. 

**MMIS: LD**

Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg \max_{f(s)} \left\{ \frac{\lambda f(s)}{\mu f(s)} (1 - e^{-\mu f(s) (t-s)}) \right\}.$$ 

On optimal path, background process may jump at most $d - 1$ times. Idea: maximize parameter of Poisson distribution. (See paper with Blom, OR Letters. Extension to general service times.)

Interesting shape of decay rate of $\mathbb{P}(\left|M^{(N)}(t)/N - \rho_t\right| \geq \varepsilon)$. 

**MMIS: LD**

Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg \max_{f(s)} \left\{ \frac{\lambda f(s)}{\mu f(s)} (1 - e^{-\mu f(s) (t-s)}) \right\}.$$ 

On optimal path, background process may jump at most $d - 1$ times. Idea: maximize parameter of Poisson distribution. (See paper with Blom, OR Letters. Extension to general service times.)

Interesting shape of decay rate of $\mathbb{P}(\left|M^{(N)}(t)/N - \rho_t\right| \geq \varepsilon)$. 

**MMIS: LD**

Intuition: a single path of $X(s)$ determines large deviations.

Naïve first thought: background process (essentially) stays in state $i$ that maximizes $\lambda_i/\mu_i$.

Wrong! Result: $X(s)$ close to path $f^*(s)$, defined by

$$\arg \max_{f(s)} \left\{ \frac{\lambda f(s)}{\mu f(s)} (1 - e^{-\mu f(s) (t-s)}) \right\}.$$ 

On optimal path, background process may jump at most $d - 1$ times. Idea: maximize parameter of Poisson distribution. (See paper with Blom, OR Letters. Extension to general service times.)

Interesting shape of decay rate of $\mathbb{P}(\left|M^{(N)}(t)/N - \rho_t\right| \geq \varepsilon)$.
Part III
MARKOV MODULATED ORNSTEIN-UHLENBECK (MMOU)
Central Limit Theorems, Large Deviations
MMOU: results

Recall: $M(\cdot)$ obeys the following SDE:

$$dM(t) = (\alpha X(t) - \gamma X(t) M(t))dt + \sigma X(t) dB(t);$$

$(B(t))_{t \geq 0}$ standard BM independent of $(X(t))_{t \geq 0}$. 

MMOU: results

Recall: \( M(\cdot) \) obeys the following SDE:

\[
dM(t) = (\alpha X(t) - \gamma X(t)M(t))dt + \sigma X(t) dB(t);
\]

\((B(t))_{t \geq 0}\) standard BM independent of \((X(t))_{t \geq 0}\).

Moments of \( M(t) \) can be found recursively, by solving non-homogeneous linear differential equations (in calculation of \( n \)-th moment, \((n - 1)\)-st moment is needed.)
MMOU: results

Recall: \( M(\cdot) \) obeys the following SDE:

\[
dM(t) = (\alpha X(t) - \gamma X(t)M(t)) dt + \sigma X(t) dB(t);
\]

\((B(t))_{t \geq 0}\) standard BM independent of \((X(t))_{t \geq 0}\).

Moments of \( M(t) \) can be found recursively, by solving non-homogeneous linear differential equations (in calculation of \( n \)-th moment, \((n - 1)\)-st moment is needed.)

Alternative: PDE for LT of \( M(t) \); recursive scheme for moments follows by standard differentiation procedure.
MMOU: results

Timescaling results: scale $\sigma^2 \mapsto N\sigma^2$, $\alpha \mapsto N\alpha$, and $Q \mapsto N^f Q$.

If $\gamma_i \equiv \gamma$, then $\text{Var} M^{(N)}(t)$ can again be found by ‘total variance’.
MMOU: results

Timescaling results: scale $\sigma^2 \mapsto N\sigma^2$, $\alpha \mapsto N\alpha$, and $Q \mapsto N^f Q$.

If $\gamma_i \equiv \gamma$, then $\text{Var} M^{(N)}(t)$ can again be found by ‘total variance’.

With similar techniques finite-dimensional CLT and weak convergence can be shown (starting of from PDE for LT of $M(t)$).

Again two regimes, depending on value of $f$. 
MMOU: results

- Specializing to the situation that $\gamma_i \equiv \gamma$ and $t \to \infty$, we obtain

$$\text{Var} M^{(N)}(\infty) = N \frac{\pi^T \sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \pi_i D_{ij}$$

$$= N \frac{\pi^T \sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \alpha^T \text{diag}\{\pi\} D\alpha.$$
MMOU: results

- Specializing to the situation that $\gamma_i \equiv \gamma$ and $t \to \infty$, we obtain

$$\text{Var} M^{(N)}(\infty) = N\frac{\pi^T\sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \pi_i D_{ij}$$

$$= N\frac{\pi^T\sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \alpha^T \text{diag}\{\pi\} D\alpha.$$ 

- Similar results for other cases.
MMOU: results

- Specializing to the situation that $\gamma_i \equiv \gamma$ and $t \to \infty$, we obtain

  $$\text{Var} M^{(N)}(\infty) = N\pi^T \sigma^2 \frac{2}{2\gamma} + \frac{N^{2-f}}{\gamma} \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \pi_i D_{ij}$$

  $$= N\pi^T \sigma^2 \frac{2}{2\gamma} + \frac{N^{2-f}}{\gamma} \alpha^T \text{diag}\{\pi\} D\alpha.$$  

- Similar results for other cases.

- **Dichotomy**: for $f > 1$ the variance is essentially linear in $N$, while for $f < 1$ it behaves superlinearly (more specifically, proportionally to $N^{2-f}$).
**MMOU: results**

- Specializing to the situation that $\gamma_i \equiv \gamma$ and $t \to \infty$, we obtain

\[
\text{Var} M^{(N)}(\infty) = N\frac{\pi^T \sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \sum_{i=1}^{d} \sum_{j=1}^{d} \alpha_i \alpha_j \pi_i D_{ij}
\]

\[
= N\frac{\pi^T \sigma^2}{2\gamma} + \frac{N^{2-f}}{\gamma} \alpha^T \text{diag}\{\pi\} D\alpha.
\]

- Similar results for other cases.

- **Dichotomy**: for $f > 1$ the variance is essentially linear in $N$, while for $f < 1$ it behaves superlinearly (more specifically, proportionally to $N^{2-f}$).

- **CLT** can be derived. Future work: weak convergence to OU process with appropriate parameters.
MMOU: Large deviations

Two regimes!

- First regime: $\alpha \mapsto N\alpha$, $\sigma^2 \mapsto N\sigma^2$, $Q \mapsto N^f Q$ with $f > 1$. 

Idea: Markov chain moves faster than OU processes. Hence:

we see effectively OU with parameters

$N_\alpha \propto N \pi_T$, $N\sigma^2 \propto N \pi_T\sigma^2$, $\gamma \propto \pi_T\gamma$.

$\lim_{N \to \infty} \frac{1}{N\log P(M(N)(t) \geq Na)} = -\frac{1}{2} (a - m_{\infty}(t))^2 s_{\infty}(t),$ 

where $m_{\infty}(t) = \alpha_{\infty} \gamma_{\infty} (1 - e^{-\gamma_{\infty} t})$, 

$s_{\infty}(t) = \sigma_{\infty}^2 \gamma_{\infty} (1 - e^{-2\gamma_{\infty} t}).$
MMOU: Large deviations

Two regimes!

▶ First regime: \( \alpha \mapsto N\alpha, \sigma^2 \mapsto N\sigma^2, Q \mapsto N^f Q \) with \( f > 1 \).

▶ Idea: Markov chain moves faster than OU processes. Hence: we see effectively OU with parameters \( N\alpha_\infty := N\pi^T\alpha, N\sigma^2_\infty := N\pi^T\sigma^2, \gamma_\infty := \pi^T\gamma \).
MMOU: Large deviations

Two regimes!

- First regime: $\alpha \mapsto N\alpha$, $\sigma^2 \mapsto N\sigma^2$, $Q \mapsto N^f Q$ with $f > 1$.
- Idea: Markov chain moves faster than OU processes. Hence: we see effectively OU with parameters $N\alpha_\infty := N\pi^T\alpha$, $N\sigma^2_\infty := N\pi^T\sigma^2$, $\gamma_\infty := \pi^T\gamma$.

$$
\lim_{N \to \infty} \frac{1}{N} \log \mathbb{P}(M^{(N)}(t) \geq Na) = -\frac{1}{2} \frac{(a - M_\infty(t))^2}{S_\infty(t)},
$$

where

$$
M_\infty(t) = \frac{\alpha_\infty}{\gamma_\infty} (1 - e^{-\gamma_\infty t}),
$$

$$
S_\infty(t) = \frac{\sigma^2_\infty}{2\gamma_\infty} (1 - e^{-2\gamma_\infty t}).
$$
MMOU: Large deviations

Proof technique:

▶ Construct lower bound by considering specific scenario.
MMOU: Large deviations

Proof technique:

- Construct lower bound by considering specific scenario.
- Split interval in subintervals of length $t/\sqrt{N}$. 

- Within each interval consider scenario that background process is close to $\pi$, viz. in $\delta$-environment.
MMOU: Large deviations

Proof technique:

- Construct lower bound by considering specific scenario.
- Split interval in subintervals of length $t/\sqrt{N}$.
- Within each interval consider scenario that background process is close to $\pi$, viz. in $\delta$-environment.
MMOU: Large deviations

Proof technique:

- Construct lower bound by considering specific scenario.
- Split interval in subintervals of length $t/\sqrt{N}$.
- Within each interval consider scenario that background process is close to $\pi$, viz. in $\delta$-environment.
- Find lower bound on mean and upper bound on variance of the Normally distribution $M^{(N)}(t)$. 

\[ \delta \downarrow 0 \]
MMOU: Large deviations

Proof technique:

- Construct lower bound by considering specific scenario.
- Split interval in subintervals of length $t/\sqrt{N}$.
- Within each interval consider scenario that background process is close to $\pi$, viz. in $\delta$-environment.
- Find lower bound on mean and upper bound on variance of the Normally distribution $M^{(N)}(t)$.
- Let $\delta \downarrow 0$. 
 MMOU: Large deviations

Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
MMOU: Large deviations

Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
- Split interval in $N^{\varepsilon/2}$ subintervals of length $t/N^{\varepsilon/2}$.
Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
- Split interval in $N^{\varepsilon/2}$ subintervals of length $t/N^{\varepsilon/2}$.
- For any event $E$,

\[
P(M^{(N)}(t) \geq Na) \leq P(M^{(N)}(t) \geq Na, E) + P(E^c).
\]

Let $E$ be the event of being close to $\pi$ (i.e., $\delta$-environment).
MMOU: Large deviations

Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
- Split interval in $N^{\varepsilon/2}$ subintervals of length $t/N^{\varepsilon/2}$.
- For any event $E$,

$$\mathbb{P}(M^{(N)}(t) \geq Na) \leq \mathbb{P}(M^{(N)}(t) \geq Na, E) + \mathbb{P}(E^c).$$

Let $E$ be the event of being close to $\pi$ (i.e., $\delta$-environment).
- Second term decays superexponentially.
MMOU: Large deviations

Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
- Split interval in $N^{\varepsilon/2}$ subintervals of length $t/N^{\varepsilon/2}$.
- For any event $E$,

$$\mathbb{P}(M^{(N)}(t) \geq Na) \leq \mathbb{P}(M^{(N)}(t) \geq Na, E) + \mathbb{P}(E^c).$$

Let $E$ be the event of being close to $\pi$ (i.e., $\delta$-environment).
- Second term decays superexponentially.
- Find upper bound on mean and lower bound on variance of the Normally distribution $M^{(N)}(t)$ on $E$. 
Proof technique, ctd.:

- Construct upper bound by showing all other scenarios are less likely, as follows.
- Split interval in $N^{\varepsilon/2}$ subintervals of length $t/N^{\varepsilon/2}$.
- For any event $E$,

$$
P(M^{(N)}(t) \geq Na) \leq P(M^{(N)}(t) \geq Na, E) + P(E^c).$$

Let $E$ be the event of being close to $\pi$ (i.e., $\delta$-environment).
- Second term decays superexponentially.
- Find upper bound on mean and lower bound on variance of the Normally distribution $M^{(N)}(t)$ on $E$.
- Let $\delta \downarrow 0$. 
MMOU: Large deviations

Two regimes!

- Second regime: $\alpha \mapsto N\alpha$, $\sigma^2 \mapsto N\sigma^2$, $Q$ unchanged.
MMOU: Large deviations

Two regimes!

- Second regime: \( \alpha \mapsto N\alpha, \sigma^2 \mapsto N\sigma^2, Q \) unchanged.
- A *single* path \( f(t) \) of \( X(t) \) determines asymptotics.
MMOU: Large deviations

Two regimes!

- Second regime: $\alpha \mapsto N\alpha$, $\sigma^2 \mapsto N\sigma^2$, $Q$ unchanged.
- A single path $f(t)$ of $X(t)$ determines asymptotics.
- $M_f(t) = \mathbb{E}(M(t) \mid X = f)$ and $s_f(t) = \text{Var}(M(t) \mid X = f)$

$$\min_{f : f(t) \in \{1, \ldots, d\}} \frac{(a - M_f(t))^2}{s_f(t)}.$$
MMOU: Large deviations

Two regimes!

- Second regime: $\alpha \mapsto N\alpha$, $\sigma^2 \mapsto N\sigma^2$, $Q$ unchanged.
- A single path $f(t)$ of $X(t)$ determines asymptotics.
- $M_f(t) = \mathbb{E}(M(t) \mid X = f)$ and $S_f(t) = \text{Var}(M(t) \mid X = f)$

$$
\min_{f: f(t) \in \{1, \ldots, d\}} \frac{(a - M_f(t))^2}{S_f(t)}.
$$

- Which path optimizes this decay rate?
MMOU: Large deviations

Goal: estimate $\mathbb{P}(M(t) \geq a)$ for large $a$ (rare event).
MMOU: Large deviations

Goal: estimate $P(\mathcal{M}(t) \geq a)$ for large $a$ (rare event).

A few thoughts on rare-event simulation by importance sampling:
MMOU: Large deviations

Goal: estimate $\mathbb{P}(M(t) \geq a)$ for large $a$ (rare event).

A few thoughts on rare-event simulation by importance sampling:

- Two sources of randomness: in background process $X(\cdot)$ and in individual OU processes $U_i(\cdot)$. 
MMOU: Large deviations

Goal: estimate $\mathbb{P}(M(t) \geq a)$ for large $a$ (rare event).

A few thoughts on rare-event simulation by importance sampling:

- Two sources of randomness: in background process $X(\cdot)$ and in individual OU processes $U_i(\cdot)$.
- Change-of-measure can be constructed?
MMOU: Large deviations

Goal: estimate $\mathbb{P}(M(t) \geq a)$ for large $a$ (rare event).

A few thoughts on rare-event simulation by importance sampling:

- Two sources of randomness: in background process $X(\cdot)$ and in individual OU processes $U_i(\cdot)$.
- Change-of-measure can be constructed?
- ‘Hybrid’ idea: sample background process, and then compute probability.
Part IV
MULTIPLE COUPLED INFINITE-SERVER QUEUES (MMIS)
COUPLED INFINITE-SERVER QUEUES

Idea: single background process modulates multiple queues (cf. open problem Peter Taylor).

Example: classical Markov fluid model. When $X(t) = i$ the first queue ‘grows’ at a deterministic rate $r_i$ (regulated at 0: content cannot become negative), the second at a deterministic rate $s_i$ (regulated at 0).
COUPLED INFINITE-SERVER QUEUES

Idea: single background process modulates multiple queues (cf. open problem Peter Taylor).

Example: classical Markov fluid model. When $X(t) = i$ the first queue ‘grows’ at a deterministic rate $r_i$ (regulated at 0: content cannot become negative), the second at a deterministic rate $s_i$ (regulated at 0). As a result: queues are correlated.
COUPLED INFINITE-SERVER QUEUES

Idea: single background process modulates multiple queues (cf. open problem Peter Taylor).

Example: classical Markov fluid model. When \( X(t) = i \) the first queue ‘grows’ at a deterministic rate \( r_i \) (regulated at 0: content cannot become negative), the second at a deterministic rate \( s_i \) (regulated at 0). As a result: queues are correlated.

How to solve joint distribution? (I know solution only when \( r_i \geq s_i \), for \( i = 1, \ldots, d \).)
COUPLED INFINITE-SERVER QUEUES

Things do work out in an infinite-server context, though. Two-dimensional PGF satisfies (for stationary version of Model I):

\[
p(w, z) Q + (w - 1) \left( p(w, z) \text{diag}\{\lambda_1\} - \frac{\partial p}{\partial w} \text{diag}\{\mu_1\} \right) \\
+ (z - 1) \left( p(w, z) \text{diag}\{\lambda_2\} - \frac{\partial p}{\partial z} \text{diag}\{\mu_2\} \right) = 0.
\]
COUPLED INFINITE-SERVER QUEUES

Things do work out in an infinite-server context, though. Two-dimensional PGF satisfies (for stationary version of Model I):

\[ p(w, z) Q + (w - 1) \left( p(w, z) \text{diag}\{\lambda_1\} - \frac{\partial p}{\partial w} \text{diag}\{\mu_1\} \right) \]
\[ + (z - 1) \left( p(w, z) \text{diag}\{\lambda_2\} - \frac{\partial p}{\partial z} \text{diag}\{\mu_2\} \right) = 0. \]

Similar systems for transient distributions/Model II.
COUPLED INFINITE-SERVER QUEUES

Leads to:
COUPLED INFINITE-SERVER QUEUES

Leads to:

- two-dimensional recursions for mixed moments,
COUPLED INFINITE-SERVER QUEUES

Leads to:

- two-dimensional recursions for mixed moments,
- CLT under timescalings.
COUPLED INFINITE-SERVER QUEUES

Leads to:

- two-dimensional recursions for mixed moments,
- CLT under timescalings.
- Same can be done for MMOU.
COUPLED INFINITE-SERVER QUEUES

Relevant model: different components react to the same ‘outer world’ $\mapsto$ correlation between components.
COUPLED INFINITE-SERVER QUEUES

Relevant model: different components react to the same ‘outer world’ $\mapsto$ correlation between components.
COUPLED INFINITE-SERVER QUEUES

Relevant model: different components react to the same ‘outer world’ → correlation between components.

- For instance: interest rates can be modeled as MMOU processes (regime switching: ‘good economy’ and ‘bad economy’).
  They react in a similar way to the background process. Model enables systematic study of effect of correlation.
COUPLED INFINITE-SERVER QUEUES

Relevant model: different components react to the same ‘outer world’ \(\leftrightarrow\) correlation between components.

- For instance: interest rates can be modeled as MMOU processes (regime switching: ‘good economy’ and ‘bad economy’). They react in a similar way to the background process. Model enables systematic study of effect of correlation.

- Other example: way to correlate stock prices. Gives handle on how to exploit potential negative correlations to optimize portfolio.
COUPLED INFINITE-SERVER QUEUES

Relevant model: different components react to the same ‘outer world’ → correlation between components.

- For instance: interest rates can be modeled as MMOU processes (regime switching: ‘good economy’ and ‘bad economy’). They react in a similar way to the background process. Model enables systematic study of effect of correlation.

- Other example: way to correlate stock prices. Gives handle on how to exploit potential negative correlations to optimize portfolio.

- Also applicable in context of communication networks.
CONCLUSIONS

Many open problems:
CONCLUSIONS

Many open problems:

- one could ‘Markov-modulate’ any SDE — e.g. Cox-Ingersoll-Ross;
CONCLUSIONS

Many open problems:

- one could ‘Markov-modulate’ any SDE — e.g. Cox-Ingersoll-Ross;
- importance sampling for estimating rare-event probabilities, exact asymptotics;
CONCLUSIONS

Many open problems:

- one could ‘Markov-modulate’ any SDE — e.g. Cox-Ingersoll-Ross;
- importance sampling for estimating rare-event probabilities, *exact* asymptotics;
- solve variational problem associated with LD for MMOU;
CONCLUSIONS

Many open problems:

- one could ‘Markov-modulate’ any SDE — e.g. Cox-Ingersoll-Ross;
- importance sampling for estimating rare-event probabilities, exact asymptotics;
- solve variational problem associated with LD for MMOU;
- distribution of running maximum of MMIS and MMOU.
CONCLUSIONS, ctd.

- MMIS and MMOU models allow fairly explicit analysis;
CONCLUSIONS, ctd.

- MMIS and MMOU models allow fairly explicit analysis;
- various asymptotic regimes can be explored;
CONCLUSIONS, ctd.

▶ MMIS and MMOU models allow fairly explicit analysis;
▶ various asymptotic regimes can be explored;
▶ ... and there is still a lot of work to be done.