Refining Mean-field Approximations by Dynamic State Truncation

FRANCESCA RANDONE, IMT School For Advanced Studies Lucca, Italy
LUCA BORTOLUSSI, Università degli Studi di Trieste, Italy
MIRCO TRIBASTONE, IMT School For Advanced Studies Lucca, Italy

Mean-field models are an established method to analyze large stochastic systems with N interacting objects by means of simple deterministic equations that are asymptotically correct when N tends to infinity. For finite N, mean-field equations provide an approximation whose accuracy is model- and parameter-dependent. Recent research has focused on refining the approximation by computing suitable quantities associated with expansions of order $1/N$ and $1/N^2$ to the mean-field equation. In this paper we present a new method for refining mean-field approximations. It couples the master equation governing the evolution of the probability distribution of a truncation of the original state space with a mean-field approximation of a time-inhomogeneous population process that dynamically shifts the truncation across the whole state space. We provide a result of asymptotic correctness in the limit when the truncation covers the state space; for finite truncations, the equations give a correction of the mean-field approximation. We apply our method to examples from the literature to show that, even with modest truncations, it is effective in models that cannot be refined using existing techniques due to non-differentiable drifts, and that it can outperform the state of the art in challenging models that cause instability due orbit cycles in their mean-field equations.

CCS Concepts: • **General and reference** → Performance; • **Mathematics of computing** → **Stochastic processes**; **Ordinary differential equations**.

Additional Key Words and Phrases: mean-field models, Markov population processes, mean estimation, state-space truncation

ACM Reference Format:
Francesca Randone, Luca Bortolussi, and Mirco Tribastone. 2021. Refining Mean-field Approximations by Dynamic State Truncation. *Proc. ACM Meas. Anal. Comput. Syst.* 5, 2, Article 25 (June 2021), 30 pages. https://doi.org/10.1145/3460092

1 INTRODUCTION

Mean-field models are a well-known technique for the analysis of stochastic systems describing a population of N interacting objects [1, 20]. When N is large, the exact analysis of such models is generally prohibitive: since closed-form solutions are available only in special cases, one must resort to computationally demanding numerical simulations to cope with state spaces that grow exponentially with m^N in the worst case, where $m \ll N$ is the cardinality of the state space of the individual object. Mean-field theory, instead, provides a simple system of differential (or difference) equations of size m with guarantees of convergence of the stochastic model as N goes to infinity

Authors' addresses: Francesca Randone, Mirco Tribastone, IMT School for Advanced Studies Lucca, Piazza San Francesco 19, 55100 Lucca, Italy; Luca Bortolussi, Department of Mathematics and Geosciences, University of Trieste, Via A. Valerio 12/1, 34127 Trieste, Italy.

2476-1249/2021/6-ART25 $15.00
https://doi.org/10.1145/3460092

Table 1. Transient average queue length (normalized by N) at time $t = 100$ (computed by simulation) and mean-field estimate for an $M/Cox/N$ queue with arrivals at rate $N\lambda$, with $\lambda = 0.75$, and a two-phase Coxian distribution with unitary service time and varying variance (V).

Scaling	$V = 5$	$V = 10$	$V = 20$
$N = 1$	6.09	8.23	10.27
$N = 5$	1.47	1.89	2.30
$N = 10$	0.97	1.10	1.25
Mean field approximation	0.75	0.75	0.75

both in the transient and in the steady-state [1, 20]. This is appealing in a variety of applications including network protocols [6], information-spreading algorithms [7], peer-to-peer networks [24], caching algorithms [14], garbage collection [37], and load balancing strategies [13, 26, 27, 34, 40].

The asymptotic results of convergence of mean-field theory provide no guarantees as to the quality of the approximation for finite N, which is model- and parameter-dependent in general. This has stimulated research into providing error bounds [21, 41, 42] and rates of convergence [8, 11, 39]. Recently, the problem of refining the mean-field approximation has attracted the attention from the performance evaluation community. Gast demonstrated that the expected value of a performance functional converges to the mean-field limit at rate $O(1/N)$ [11], both in the transient and in the stationary regime (under the assumption of a unique attractor that is exponentially locally stable). Later, Gast and Van Houdt used a Lyapunov argument to compute the constant associated with the $1/N$ term for the steady-state expectation of such functionals, effectively providing a refinement of the approximation for finite N [15]. The method is further extended with an expansion of the term $1/N^2$ for both the transient and the steady-state regimes [12].

These mean-field refinements fundamentally assume differentiability of the drift, i.e., the vector field of the limit system. This rules out their applicability to a class of queuing models for which corrections to the mean-field estimates of average queue lengths may be desirable. As a motivating example, let us consider an $M/Cox/N$ queue with arrivals at rate $N\lambda$, using a classic scaling approach (e.g., [8]). Table 1 shows the errors between the average queue length (normalized by N) and its mean-field approximation for service times distributed with a two-phase Coxian distribution with unitary mean and increasing variance, using the fitting formulae in [2]. The simulations, computed at an arbitrary time $t = 100$ for which the mean-field approximation approaches a stationary regime, show errors that decrease with N, but can be large for small N; in particular, the errors increase with the variance of the Coxian distribution for any fixed N. Furthermore, the mean-field approximation numerically shows insensitivity to the variance of the service process, being always equal to λ—a fact that has been discussed in a more general setting in [4]. We conclude that the mean-field approximation may not adequately represent the stochastic queue length dynamics for finite, small, N. However, the aforementioned mean-field refinement results cannot be used because the drift is piece-wise linear due to the presence of the minimum function in the rate transitions of the population process.[1]

In this paper we present a new method to refine average estimates of population processes with non-differentiable drifts. Although the main application lies with mean-field models, the method does not make use of the deterministic ODE equation as the limit behavior of the re-scaled Markov

[1]For instance, in the special case of exponentially distributed service times, i.e., for an $M/M/N$ queue, service occurs with rate $\mu \min(X, N)$ where X is the queue length and μ is the service rate. Further details are given in Section 6.

process when N goes to infinity; in particular, it does not assume a density-dependent process. Instead, it starts from the equation for the "true" average evolution of a population process $X(t)$,

$$\frac{d\mathbb{E}\left[X(t)\right]}{dt} = \mathbb{E}\left[f(X(t))\right],$$

where f is the drift [3, 33, 38]. When f is not linear, this cannot be directly used because the right-hand side cannot be expressed in terms of the variables occurring in the left-hand side. Hence the approximation of the average essentially consists in assuming $\mathbb{E}\left[f(X(t))\right] \approx f(\mathbb{E}\left[X(t)\right])$ in the above equation, yielding

$$\frac{d\mathbb{E}\left[X(t)\right]}{dt} = f(\mathbb{E}\left[X(t)\right]). \tag{1}$$

In the case of density-dependent models, the rescaled process $X^N(t)/N$ with drift f converges to the solution of the ODE $dx(t)/dt = f(x(t))$ as N tends to infinity, which has the same functional form as (1) but $x(t)$ is interpreted as a density/proportion of objects rather than a population. For this reason we shall still refer to (1) as the equations for the mean-field approximation.

Our proposal for refining average estimates is based on the definition of an auxiliary process $Y(t)$, depending on the original process $X(t)$, such that each state y of $Y(t)$ identifies a hyper-rectangular subset of states in which $X(t)$ is at time t. By using a suitable truncation of the state space, the transition rate matrix governing the evolution of the probability distribution of $X(t)$ on y can be expressed as a function of y alone. Specifically, we consider a truncation-and-augmentation scheme where the transition rates toward states outside a given truncation are redirected to the "closest" state on its boundary. Augmented truncations preserve the stochasticity of the transition rate matrix of the truncated process, as opposed to the classical approaches for Markov chains in discrete (e.g., [31, 32]) and continuous (e.g., [35, 36]) time; they have received attention due to theoretical guarantees on the convergence of their stationary distributions to that of the original process [17] and the possibility to provide useful lower and upper bounds on the steady-state probability of single states [22, 23, 25]. Importantly for our refinement method, with the proposed augmentation it is possible to approximate $Y(t)$—which is non-Markovian—as a time-inhomogeneous Markov population process for which we can write equations for the mean-field approximation.

Overall, this results in a family of refined approximations depending on a parameter $n \in \mathbb{N}^m$ that gives the size of the truncation, i.e., the volume of the hyper-rectangle that defines the truncated state space. For any finite n, the approximation consists in a system of ODEs where the master equation governing the transient probability distribution of the truncated process is modulated by a continuous variable representing the mean-field equations of $Y(t)$. Effectively, this leads to a *dynamic* shift of the truncation across the state space of the original process. Due to the specific choice of truncation scheme, we call our method *dynamic boundary projection* (DBP).

Theoretically, our main formal result concerns a statement of asymptotic correctness. We prove that, as $n \longrightarrow \infty$ and under the assumption of boundedness of the drift, the family of DBP approximations converge to the original population process $X(t)$. Usually, truncation approximation methods yield linear systems of ODEs [9, 19] for which component-wise or total variation convergence of the solution to the master equation can be proved using standard techniques (see, e.g. [28] and [17] for convergence over finite time intervals and stationary distributions, respectively). DBP yields a nonlinear ODE system due to the coupling of the (linear) master equation with the (usually nonlinear) drift of $X(t)$. In this case we need a stronger convergence result, namely uniform convergence in 1-norm in a suitable Banach space, and a different proof strategy, which we base on a perturbation argument for nonlinear ODEs.

From a numerical viewpoint, we apply DBP to three examples. First, we consider the aforementioned Coxian queuing system to show how the refinements can improve the mean-field approximation such that it is not insensitive to the variance of the service-time distribution. As discussed this is done on a population process that has a non-differentiable drift, hence available mean-field refinement methods are not applicable. Second, we consider a model of malware propagation used in [12] to show the instability in the computation of the refined $1/N$ and $1/N^2$ terms in the presence of orbit cycles in the mean-field approximation. Under these conditions we numerically show that DBP does not exhibit instability and can improve the mean-field approximations, while it performs similarly to the refined scheme of [12] if the mean-field model has a unique attractor. Finally, we consider a multi-class queuing system [30] with processor sharing discipline as a case study to show how the choice of the parameter n may impact on the quality of the approximation.

Paper structure. Section 2 presents the model and builds the joint process $(X(t), Y(t))$ on which DBP is based. Section 3 presents the boundary-projection truncation method. Section 4 derives the DBP equations to refine the mean-field approximation. The main theorem of asymptotic correctness is proved in Section 5. Section 6 presents the case studies, while conclusion and future works are discussed in Section 7. Proofs not provided in the main text are given in Appendix.

2 MODEL AND NOTATION

2.1 Markov Population Process

In this paper we consider a Markov population process $(X(t) \in \mathcal{S})_{t\geq 0}$ evolving on a set $\mathcal{S} \subseteq \mathbb{N}^m$. We denote by $H(\mathcal{S}) \subseteq \mathbb{R}^m$ the convex hull of $\mathcal{S}$. The initial state of $X(t)$ is denoted by $\bar{x}_0 \in \mathcal{S}$. Given a finite set of jump vectors $\mathcal{L} \subseteq \mathbb{Z}^m$ and a set of transition functions $f_l : H(\mathcal{S}) \to \mathbb{R}_{\geq 0}$, for $l \in \mathcal{L}$, $X(t)$ makes transitions at rate $f_l(x)$ from state x to state $x + l$ for each $l \in \mathcal{L}$.

Defining the *drift* f as

$$f(x) = \sum_{l \in \mathcal{L}} l f_l(x),$$

the mean-field approximation of $X(t)$ is the solution to the Cauchy problem:

$$\begin{cases} \frac{dx}{dt} = f(x(t)) \\ x(0) = \bar{x}_0. \end{cases} \tag{2}$$

From now on we will assume that the following two assumptions on the drift are verified:

(H1) there exists a constant $L > 0$ such that $\|f(x + h) - f(x)\|_1 \leq L\|h\|_1 \, \forall x, x + h \in H(\mathcal{S})$;
(H2) there exists a constant $C > 0$ such that $\|f(x)\|_1 \leq C \, \forall \, x \in H(\mathcal{S})$.

The first condition (local Lipschitz continuity) ensures the existence and uniqueness of the solution of (2). The second condition (boundedness of the drift) is required to prove uniform convergence in 1-norm of the proposed approximation to the original process (see Section 5). Observe that these conditions have to hold on the convex hull $H(\mathcal{S})$ and not only on the set of states, since in our approximation $f(x)$ will be evaluated also on continuous values of x.

Remark 1. *As introduced in Section 1, we stress that this formulation does not assume density dependence, hence, from now on each state of the Markov population process will be an unscaled population vector (rather than a vector describing proportions of objects in every local state). The mean-field equation (2) is thus interpreted as the approximate dynamics of the average populations.*

Remark 2. *The hypothesis $\mathcal{S} \subseteq \mathbb{N}^m$ can be easily extended to the case $\mathcal{S} \subseteq \mathbb{Z}^m$. However, it is more convenient to present the theory using the former hypothesis since all since all our examples will evolve on $\mathbb{N}^m$.*

2.2 Joint Process

We now define a second process $Y(t)$ depending on $X(t)$, whose value identifies the truncation in which X is at time t. We fix a bound $n \in \mathbb{N}^m$ and define for each $y \in \mathcal{S}$ the set of states:

$$\mathcal{T}_y^{(n)} = \{x_i \in \mathcal{S} : y_i \leq x_i \leq y_i + n_i \quad \forall\, i = 1, \ldots, m\}. \tag{3}$$

We say that $\mathcal{T}_y^{(n)}$ is a (hyper-rectangular) truncation of the state space indexed by y.

For every $x, y \in \mathcal{S}$ we define $\Pi^{(n)}(x, y) \in \mathbb{Z}^m$ as:

$$\Pi_i^{(n)}(x,y) = \begin{cases} x_i & \text{if } x_i < y_i \\ x_i - n_i & \text{if } x_i > y_i + n_i \\ y_i & \text{if } y_i - n_i \leq x_i \leq y_i + n_i \end{cases} \qquad \forall\, i = 1, \ldots, m.$$

Intuitively, $\Pi^{(n)}(x, y)$ returns a vector y' such that the truncation $\mathcal{T}_{y'}^{(n)}$ is the closest truncation to $\mathcal{T}_y^{(n)}$ that contains x. We use $\Pi^{(n)}(x, y)$ to build a coupled process $Y(t)$ that tracks the truncation in which $X(t)$ is at time t. We define

$$Y(t) = y_k \qquad \forall t \in [t_k, t_{k+1}),$$

where

$$y_0 = \Pi^{(n)}(\bar{x}_0, 0), \qquad y_k = \Pi^{(n)}(X(t_k), y_{k-1}), \qquad t_0 = 0, \qquad t_k = \inf\left\{t > t_{k-1} : X(t) \notin \mathcal{T}_{y_{k-1}}^{(n)}\right\}.$$

According to this definition, $Y(t) = y_k$ means that at time t the original process $X(t)$ is in a state belonging to the truncation $\mathcal{T}_{y_k}^{(n)}$; $Y(t)$ changes value whenever $X(t)$ makes a jump into a state x outside the current truncation. Observe that $Y(t)$ is uniquely defined although x may belong to different truncations: the new value of $Y(t)$ is the truncation containing the current state closest to the previous truncation.

Denoting the characteristic function by $\mathbb{I}_{\cdot}$, the joint process $(X(t), Y(t))$ is a time-homogeneous continuous-time Markov chain defined by the following transitions:

- $(X(t), Y(t))$ jumps from state (x, y) to state $(x + l, y)$ with rate $\mathbb{I}_{\{x+l \in \mathcal{T}_y^{(n)}\}} f_l(x)$ for every $l \in \mathcal{L}$;
- $(X(t), Y(t))$ jumps from state (x, y) to state $(x + l, \Pi^{(n)}(y, x + l))$ with rate $\mathbb{I}_{\{x+l \notin \mathcal{T}_y^{(n)}\}} f_l(x)$ for every $l \in \mathcal{L}$.

By definition of $\Pi^{(n)}$, the joint process evolves only on states (x, y) such that $x \in \mathcal{T}_y^{(n)}$.

The master equation for the transient joint probability distribution $P(x, y; t)$ of $(X(t), Y(t))$ is:

$$\frac{dP(x,y)}{dt} = -\left[\sum_{l \in \mathcal{L}} f_l(x)\right] P(x,y;t) + \tag{4}$$

$$+ \sum_{l \in \mathcal{L}} \left[f_l(x-l)P(x-l,y;t) + \sum_{y' : \Pi^{(n)}(x,y') = y} \mathbb{I}_{\{x \notin \mathcal{T}_{y'}^{(n)}\}} f_l(x-l) P(x-l, y'; t) \right]. \tag{5}$$

The sum in (4) accounts for the outgoing probability flux from (x, y), due to all possible jumps, whether exiting the current truncation or not. The first summand in (5) accounts for the incoming probability flux into state (x, y) due to the transitions taking place in states (x', y) belonging to the same truncation $\mathcal{T}_y^{(n)}$ (observe that if $x - l \notin \mathcal{T}_y^{(n)}$, then $P(x - l, y; t) = 0$ for all t). The second summand in (5) accounts for the incoming probability flux due to those transitions starting in states (x', y') belonging to different truncations $\mathcal{T}_{y'}^{(n)}$ but having (x, y) as target state.

Example. *In the rest of the paper we will use an $M/M/k$ queue as a running example. It is defined by the following transition classes, denoting respectively exogenous arrivals with Poisson rate λ and service with rate μ, respectively:*

$$\begin{aligned} l_1 &= +1, && \text{at rate } \lambda, \\ l_2 &= -1 && \text{at rate } \mu \min(x,k). \end{aligned}$$

Let us write the master equation of the joint process $(X(t), Y(t))$. The summand in (4) and the first summand in (5) can be rewritten directly, so we focus on the second summand in (5). For this term, the only non-zero contributions are given by those x and y' satisfying:

$$\begin{aligned} x \notin \mathcal{T}_{y'}^{(n)} &\Rightarrow (x < y) \vee (x > y' + n), \\ x - l \in \mathcal{T}_{y'}^{(n)} &\Rightarrow y' \leq x - l \leq y' + n, \\ \Pi^{(n)}(x, y') &= y. \end{aligned}$$

Applying the first two conditions with $l_1 = +1$ gives $x = y' + n + 1$, while from the last condition we get $y' = y - 1$; thus, this term appears in the master equation only when $x = y + n$. This agrees with the intuition that, when the process is in state $(y - 1 + n, y - 1)$ (the last state of the truncation indexed by $y - 1$) and a new customer arrives, it jumps to the next truncation, indexed by y.

Applying the same argument to the service transition, we get $x = y' - 1$ and $y' = y + 1$, so that this term appears in the equation only when $x = y$. That is, from state $(y + 1, y + 1)$ (the first state of the truncation indexed by $y + 1$), upon service the process jumps to the previous truncation, indexed by y.

The resulting master equation is:

$$\begin{aligned} \frac{d}{dt} P_{x,y}(t) = &- (\lambda + \mu \min(x,k)) P_{x,y}(t) + \lambda P_{x-1,y}(t) + \mu \min(x+1,k) P_{x+1,y}(t) \\ &+ \mathbb{I}_{\{x=y+n\}} \lambda P_{y+n-1,y-1}(t) + \mathbb{I}_{\{x=y\}} \mu \min(y+1,k) P_{y+1,y+1}(t). \end{aligned}$$

□

From (4)-(5) we can derive the expression for the exact mean of $X(t)$ as it follows:

$$\begin{aligned} \mathbb{E}[X(t)] &= \sum_{x \in \mathcal{S}} x P(x;t) = \sum_{x \in \mathcal{S}} \sum_{y : x \in \mathcal{T}_y^{(n)}} x P(x,y;t) \\ &= \sum_{y \in \mathcal{S}} \sum_{x \in \mathcal{T}_y^{(n)}} x P(x,y;t) \\ &= \sum_{y \in \mathcal{S}} P(y;t) \sum_{x \in \mathcal{T}_y^{(n)}} x P(x|\, y;t). \end{aligned}$$

We observe that we can express all the states in a truncation $\mathcal{T}_y^{(n)}$ shifting by y the states in $\mathcal{T}_0^{(n)}$, so that the previous expression can be rewritten as

$$\mathbb{E}[X(t)] = \sum_{y \in \mathcal{S}} P(y;t) \sum_{x \in \mathcal{T}_0^{(n)}} (x+y) P(x+y|\, y;t). \tag{6}$$

The above expression is exact. In Section 3 we consider an approximation of the inner sum by means of a truncation of the state space. Based on this, in Section 4 we study a mean-field approximation of the outer sum. We prove the asymptotic correctness of these approximations in Section 5.

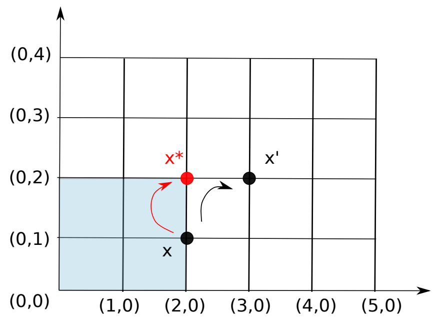

Fig. 1. Example of the application of BP on $\mathcal{T}_{(0,0)}^{(2,2)}$. A reaction from $(2,1)$ to $(3,2)$ is redirected to $(2,2)$, the state in the current truncation closest to the real target state.

3 AUGMENTED TRUNCATION APPROXIMATIONS

Let us fix a value of n. Since the inner sum of (6) considers only states in a truncation $\mathcal{T}_y^{(n)}$, here we aim to provide an approximate distribution of $X(t)$ on such subset of the state space. In order to do so we consider an augmented truncation (see, e.g., [23] and references therein). It corresponds to restricting the transition rate matrix of $X(t)$ to the states in $\mathcal{T}_y^{(n)}$ and redirecting the transitions from states $x \in \mathcal{T}_y^{(n)}$ to states $x' \notin \mathcal{T}_y^{(n)}$ to new target states $x^* \in \mathcal{T}_y^{(n)}$, with the same rates.

By doing this it is possible to define a new process $X_y^{(n)}(t)$ with transition rate matrix $Q^{(n)}(y)$, such that if $X(t)$ and $X_y^{(n)}(t)$ start from the same initial condition in $\mathcal{T}_y^{(n)}$, they will evolve identically until $t^* = \inf\{t > 0 : X(t) \notin \mathcal{T}_y^{(n)}\}$. At time t^*, $X(t)$ jumps outside $\mathcal{T}_y^{(n)}$, while $X_y^{(n)}$ will perform a transition with identical rate to a state still in $\mathcal{T}_y^{(n)}$.

Observe that the transitions that need to be redirected are those starting in states "on the border" of a truncation. This leads us to define the following sets

$$\partial\mathcal{T}_y^{(l,n)} = \left\{x \in \mathcal{T}_y^{(n)} : x + l \notin \mathcal{T}_y^{(n)}\right\}, \text{ for } l \in \mathcal{L},$$
$$\partial\mathcal{T}_y^{(n)} = \bigcup_{l\in\mathcal{L}} \mathcal{T}_y^{(l,n)} = \left\{x \in \mathcal{T}_y^{(n)} : \exists\, l \in \mathcal{L} \text{ s.t. } x + l \notin \mathcal{T}_l^{(n)}\right\}$$

for which it is immediate to show that:

$$\partial\mathcal{T}_{y+w}^{(l,n)} = \partial\mathcal{T}_y^{(l,n)} + w = \left\{x + w : x \in \partial\mathcal{T}_y^{(l,n)}\right\}.$$

We introduce a specific augmented truncation that we call *boundary projection* (BP), in which every jump from $x \in \partial\mathcal{T}_y^{(l,n)}$ is redirected with same rate to a state x^* defined as:

$$x_i^* = \begin{cases} \min(y_i + n_i, x_i') & \text{if } x_i' > x_i \\ \max(y_i, x_i') & \text{if } x_i' < x_i \\ x_i & \text{if } x_i' = x_i. \end{cases} \tag{7}$$

Essentially, x^* is the projection of the target state on the boundary of the current truncation; in other words, it is the state in the current truncation closest to the real target state (see Figure 1).

After performing the augmentation, for each state $x \in \mathcal{T}_y^{(n)}$ we have a set of jump vectors $\tilde{\mathcal{L}}^{(n)}(x)$ such that for every $l \in \mathcal{L}$ we can now define a new vector $\tilde{l}^{(n)}(x)$ given by:

$$\tilde{l}^{(n)}(x) = \begin{cases} l & \text{if } x \notin \partial\mathcal{T}_y^{(l,n)}, \\ x^* - x & \text{if } x \in \partial\mathcal{T}_y^{(l,n)}, \end{cases}$$

where x^* is the target state in which the transition $x \longrightarrow x + l$ has been redirected and the associated transition rates are $f_{\tilde{l}^{(n)}(x)}(x) = f_l(x)$.

With this choice, the transition rate matrices $Q^{(n)}(y)$ have the same functional form for every y. This will be derived by the following invariance property.

Proposition 3.1. *Assume that $x \in \partial\mathcal{T}_y^{(l,n)}$ and for $X_y^{(n)}$ transition l is redirected to state $x^* \in \mathcal{T}_y^{(n)}$, defined by BP as in (7). Then, for any $w \in \mathbb{Z}^m$, for the process $X_{y+w}^{(n)}$, the transition from $x + w \in \partial\mathcal{T}_{y+w}^{(l,n)}$ is redirected to $x^* + w \in \mathcal{T}_{y+w}^{(n)}$.*

We can now explicitly derive the master equation for $X_y^{(n)}$. Observe that for each truncation the total number of states is $\mathcal{N}(n) = \prod_{i=1}^{m}(n_i + 1)$. For a fixed y, using the definition in (7) we obtain the rate transition matrix $Q^{(n)}(y)$ where each component $[Q_y^{(n)}]_{x,x'}$ is given by:

$$[Q^{(n)}(y)]_{x,x'} = \begin{cases} \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x'+\tilde{l}^{(n)}(x')=x\}} f_l(x') & \text{if } x \neq x', \\ -\sum_{l\in\mathcal{L}} \mathbb{I}_{\{\tilde{l}^{(n)}(x)\neq 0\}} f_l(x) & \text{if } x = x', \end{cases} \qquad \text{for } x, x' \in \mathcal{T}_y^{(n)}.$$

Observe that if a transition starting in x is redirected to x itself, we remove this transition from the approximated process $X_y^{(n)}$. Since Proposition 3.1 assures the invariance by translation of the target states, and all the states in a given truncation $\mathcal{T}_y^{(n)}$ can be written shifting by y the states in $\mathcal{T}_0^{(n)}$, we can rewrite the previous matrix as follows:

$$[Q^{(n)}(y)]_{x,x'} = \begin{cases} \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x'+\tilde{l}^{(n)}(x')=x\}} f_l(x'+y) & \text{if } x \neq x' \\ -\sum_{l\in\mathcal{L}} \mathbb{I}_{\{\tilde{l}^{(n)}(x)\neq 0\}} f_l(x+y) & \text{if } x = x' \end{cases} \qquad \text{for } x, x' \in \mathcal{T}_0^{(n)}. \tag{8}$$

We can see from this expression that the functional form of the transition rate matrix remains the same as y varies. Then, the resulting master equation for BP can be written as:

$$\frac{dP_y^{(n)}}{dt} = Q^{(n)}(y)P_y^{(n)}(\,\cdot\,;t) \tag{9}$$

where $P_y^{(n)}(\,\cdot\,;t)$ is an $\mathcal{N}(n)$-dimensional vector indexed by the states in $\mathcal{T}_0^{(n)}$ and each component $P_y^{(n)}(x;t)$ gives the probability of $X_y^{(n)}$ being in the state $x + y$.

We can now approximate the inner sum in (6) using $P_y^{(n)}(\cdot;t)$ as an approximation for $P(\cdot+y|\,y;t)$. This is justified by the fact that, if the original process $X(t)$ never exits the truncation $\mathcal{T}_y^{(n)}$ for all $0 \leq t \leq T$, where T is a given finite time horizon, then $P_y^{(n)}(\,\cdot\,;t) = P(\,\cdot\, + y|\,y;t)$ for all $0 \leq t \leq T$. The resulting approximation is:

$$\mathbb{E}[X(t)] \approx \sum_{y\in\mathcal{S}} P(y;t) \sum_{x\in\mathcal{T}_0^{(n)}} (x+y)P_y^{(n)}(x;t). \tag{10}$$

Example. *Let us now apply BP to the running example of $M/M/k$ queue in $\mathcal{T}_y^{(n)}$. We have $\partial\mathcal{T}_y^{(-1,n)} = \{y\}$ and $\partial\mathcal{T}_y^{(+1,n)} = \{y+n\}$, thus we need to redirect the transitions $y \longrightarrow y-1$ and $y+n \longrightarrow y+n+1$. According to (7), they will have new target states y and $y+n$, respectively. Observe* *that, since in both cases the transitions are redirected to the starting states, we will remove these transitions from the approximated process.*

The master equation of the approximated process on $\mathcal{T}_y^{(n)}$ is then given by:

$$\frac{dP_y^{(n)}(x)}{dt} = \begin{cases} -\lambda P_y^{(n)}(0;t) + \mu \min(1+y,k) P_y^{(n)}(1;t) & \text{if } x = 0 \\ -(\lambda + \mu \min(x+y,k)) P_y^{(n)}(x;t) + \lambda P_y^{(n)}(x-1;t) + \\ \qquad + \mu \min(x+1+y,k) P_y^{(n)}(x+1;t) & \text{if } x = 1,\dots,n-1 \\ -\mu \min(n+y,k) P_y^{(n)}(n;t) + \lambda P_y^{(n)}(n-1;t) & \text{if } x = n \end{cases}$$

where $P_y^{(n)}(x;t)$ is interpreted as the probability of $X_y^{(n)}$ being in the state $x+y$. Observe that this is equivalent to an $M/M/k/n$ queue in which there are always at least y jobs in the queue. Moreover, choosing a different value of y does not change the functional form of the vector field $Q^{(n)}(y)$. □

4 DYNAMIC BOUNDARY PROJECTION

To approximate the outer sum in (10), from (4)-(5) we can derive the equations for the evolution of marginal probability distribution $P(y;t) = \sum_{x \in \mathcal{S}} P(x,y;t)$ of the joint process.

After some simple manipulations, we get:

$$\frac{dP(y)}{dt} = -\sum_{l \in \mathcal{L}} \sum_{x \in \mathcal{S}} \mathbb{I}_{\{x+l \notin \mathcal{T}_y^{(n)}\}} f_l(x) P(x,y;t) + \sum_{l \in \mathcal{L}} \sum_{x} \sum_{y' : \Pi^{(n)}(x+l,y')=y} \mathbb{I}_{\{x+l \notin \mathcal{T}_{y'}^{(n)}\}} f_l(x) P(x,y';t), \tag{11}$$

where the first term in the right-hand side describes the outgoing probability flux from $\mathcal{T}_y^{(n)}$, due to transitions taking place in states $x \in \partial\mathcal{T}_y^{(n)}$; the second term instead describes the incoming probability flux into $\mathcal{T}_y^{(n)}$ due to transitions from states $x \in \partial\mathcal{T}_{y'}^{(n)}$ such that $x + l \in \mathcal{T}_y^{(n)}$ for some l. Observe that, since in general $Y(t)$ is not Markovian, it is not possible to express the equations for $Y(t)$ in a closed form independent from $X(t)$.

Similarly to what done in (6), we now express all the states in a truncation $\mathcal{T}_y^{(n)}$ shifting by y the states in $\mathcal{T}_0^{(n)}$. For a fixed y, the possible values of y' such that $\Pi^{(n)}(x+l,y') = y$ can be obtained shifting by y the values of y' such that $\Pi^{(n)}(x+l,y') = 0$. This follows from the following property.

Proposition 4.1. $\Pi^{(n)}(x,y) = \Pi^{(n)}(x-y,0) + y$ *for all* $x, y \in \mathcal{S}$.

Thus we have $\Pi^{(n)}(x+y'+l,y') = y$ if and only if $\Pi^{(n)}(x+l,0) = y - y'$; therefore, for a fixed y, the values taken by y' are given exactly by $y - \Pi^{(n)}(x+l,0)$ for $l \in \mathcal{L}$ and all $x \in \partial\mathcal{T}_0^{(l,n)}$. This leads us to define the function $\mathcal{Y}^{(l,n)}(x) = \Pi^{(n)}(x+l,0)$ for all $l \in \mathcal{L}, x \in \partial\mathcal{T}_0^{(l,n)}$.

Example. *For an $M/M/k$ queue and a fixed y, we want to find the values of y' for which $\Pi^{(n)}(x+l,y') = y$, i.e., we want to find the truncations $\mathcal{T}_{y'}^{(n)}$ from which we can transition to $\mathcal{T}_y^{(n)}$. As discussed, this happens when an arrival takes place in $(n+y-1, y-1)$ and when a service takes place in $(y+1, y+1)$; thus, the possible values of y' are exactly $y-1$ and $y+1$. We can obtain this values also in the following way:*

- *for $l = -1$, $\partial\mathcal{T}_0^{(-1,n)} = \{0\}$, so $y' = y - \mathcal{Y}^{(-1,n)}(0) = \Pi^{(n)}(-1,0) = -1$;*
- *for $l = +1$, $\partial\mathcal{T}_0^{(+1,n)} = \{n\}$, so $y' = y - \mathcal{Y}^{(+1,n)}(n) = \Pi^{(n)}(n+1,0) = 1$.*

□

We now replace the joint probabilities with the conditionals and rewrite (11) as:

$$\frac{dP(y)}{dt} = -\sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} f_l(x+y)P(x+y|\, y;t)P(y;t)+$$
$$+\sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} f_l(x{+}y{-}\mathcal{Y}^{(l,n)}(x))P(x{+}y{-}\mathcal{Y}^{(l,n)}(x)|\, y{-}\mathcal{Y}^{(l,n)}(x);t)P(y{-}\mathcal{Y}^{(l,n)}(x);t).$$

Again we can approximate the conditional probabilities $P(\,\cdot+y|\, y;t)$ with $P_y^{(n)}(\,\cdot\,;t)$. The resulting approximated equation for $P(y,t)$ is:

$$\frac{dP(y)}{dt} \approx \sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} f_l(x+y)P_y^{(n)}(x;t)P(y;t)+$$
$$+\sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} f_l(x+y-\mathcal{Y}^{(l,n)}(x))P^{(n)}_{y-\mathcal{Y}^{(l,n)}(x)}(x;t)P(y-\mathcal{Y}^{(l,n)}(x);t). \quad (12)$$

By doing this, we are implicitly approximating $Y(t)$ as a time-inhomogeneous Markov process $Y^{(n)}$. Indeed, while the (exact) conditional probabilities $P(\,\cdot\,+y|\, y;t)$ depend on $X(t)$ and cannot be computed knowing only the state of $Y(t)$ at a previous time instant, the approximated probabilities $P_y^{(n)}(\,\cdot\,;t)$ depend on the value of $Y^{(n)}(t)$ alone. In light of this, (12) can be interpreted as the master equation of $Y^{(n)}$, in which the first summand describes the outgoing probability flow from a state y and the second summand represents the incoming probability flow to state y from states $y-\mathcal{Y}^{(l,n)}(x)$. In particular, $Y^{(n)}(t)$ is a Markov population process with jump vectors $\mathcal{Y}^{(l,n)}(x)$ and transition functions $f_l(x+y)P_y^{(n)}(x;t)$ for all $l\in\mathcal{L}$ and $x\in\partial\mathcal{T}_0^{(l,n)}$, where $P_y^{(n)}(x;t)$ is a function of y, since it is the solution of system (9), with transition matrix $Q^{(n)}(y)$. Observe that the transition functions are time-dependent since $P_y^{(n)}(\,\cdot\,;t)$ is. Thus we can write the mean-field approximation of $Y^{(n)}$, which gives:

$$\frac{dY^{(n)}}{dt} = \sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} \mathcal{Y}^{(l,n)}(x)f_l(x+Y^{(n)}(t))P^{(n)}_{Y^{(n)}}(x;t) \quad (13)$$

We can now approximate the distribution $P(y;t)$ appearing in (10) with a delta distribution peaked on the state $Y^{(n)}$ solution of (13). The result is the following approximation of the mean:

$$\mathbb{E}[X(t)] \approx \sum_{x\in\mathcal{T}_0^{(n)}} (x+Y^{(n)}(t))P^{(n)}_{Y^{(n)}}(x;t). \quad (14)$$

The quantities appearing in this approximation are governed by the following set of equations, which we call the *dynamic boundary projection* (DBP) of $X(t)$:

$$\frac{dY^{(n)}}{dt} = \sum_{l\in\mathcal{L}}\sum_{x\in\partial\mathcal{T}_0^{(l,n)}} \mathcal{Y}^{(l,n)}(x)f_l(x+Y^{(n)}(t))P^{(n)}(x;t) \quad (15)$$

$$\frac{dP^{(n)}}{dt} = Q^{(n)}(Y^{(n)})P^{(n)}(\,\cdot\,,t). \quad (16)$$

Observe that we removed the subscript $Y^{(n)}$ from $P^{(n)}$ since from now on we will always consider a single truncated master equation with time-varying transition rate matrix $Q^{(n)}(Y^{(n)})$.

The initial conditions for the DBP equations are given by those of the joint process $(X(t), Y(t))$ that have been defined as $(X(0), Y(0)) = (\bar{x}_0, \Pi^{(n)}(0, \bar{x}_0)) = (\bar{x}_0, \bar{y}_0)$. Since $\bar{x}_0 \in \mathcal{T}^{(n)}_{\bar{y}_0}$ we can rewrite $\bar{x}_0$ as $\bar{x}^*_0 + \bar{y}_0$ with $\bar{x}^*_0 \in \mathcal{T}^{(n)}_0$. Then we set the initial conditions for (15)-(16) as:

$$Y^{(n)}(0) = \bar{y}_0 \qquad\qquad P^{(n)}(x; 0) = \begin{cases} 1 & \text{if } x = \bar{x}^*_0, \\ 0 & \text{else.} \end{cases} \tag{17}$$

The DBP equations can be interpreted in the following way. We are considering all possible augmented truncation approximations of X on the sets $(\mathcal{T}^{(n)}_y : y \in \mathcal{S})$. For each of these approximations, (16) represents the associated master equation, which depends on the considered subset through the parameter $Y^{(n)}$. The value of such parameter evolves according to (15) and can be seen as a continuous approximation of the "most probable" truncation on which X is evolving at time t. Observe that due to the mean-field approximation we are now considering continuous values of $Y^{(n)}$. Moreover we can now interpret the r.h.s. of (14) as if $P^{(n)}(x; t)$ gives the probability at time t of a state $x + Y^{(n)}(t)$.

Example. *For the $M/M/k$ queue, applying DBP gives the following:*

$$\frac{dP^{(n)}(x)}{dt} = \begin{cases} -\lambda P^{(n)}(0; t) + \mu \min(1 + Y^{(n)}(t), k) P^{(n)}(1; t) & \text{if } x = 0 \\ -(\lambda + \mu \min(x + Y^{(n)}(t), k)) P^{(n)}(x; t) + \lambda P^{(n)}(x - 1; t) + \\ \qquad + \mu \min(x + 1 + Y^{(n)}(t), k) P^{(n)}(x + 1; t) & \text{if } x = 1, \dots, n - 1 \\ -\mu \min(n + Y^{(n)}(t), k) P^{(n)}(n; t) + \lambda P^{(n)}(n - 1; t) & \text{if } x = n \end{cases}$$

$$\frac{dY^{(n)}}{dt} = -\mu \min(Y^{(n)}(t), k) P^{(n)}(0; t) + \lambda P^{(n)}(n; t).$$

The approximate the mean evolution of X is thus given by:

$$\mathbb{E}[X(t)] \approx Y^{(n)}(t) + \sum_{i=0}^{n} i P^{(n)}(i; t).$$

□

To highlight the relation between DBP, the mean-field approximation of $X(t)$ and its full master equation for the transient probability distribution, we consider the two degenerate cases associated to $n = 0$ and $n = \infty$. According to the definition in (3), these correspond to a single-state truncation and an infinite truncation covering the whole state space, respectively.

Proposition 4.2. *The following two propositions hold:*

- *For $n = 0$, the DBP equations* (15)-(16) *yield the mean-field approximation of $X(t)$.*
- *For $n = \infty$, the DBP equations* (15)-(16) *yield the master equation of $X(t)$.*

Observe that, in particular, the second result suggests the convergence for the solutions of the DBP equations proved in the next section.

5 CONVERGENCE

Consider the DBP equations in (15)-(16) with initial condition as in (17). We denote the solution of the associated Cauchy problem with the $(m + \mathcal{N}(n))$-dimensional vector $Z^{(n)} = [Y^{(n)}, P^{(n)}]^T$. Recall that by Proposition 4.2 we know that $Z^\infty = [0, P^\infty]^T$ is the solution to the master equation of the original process $X(t)$.

For a fixed $T > 0$ and for every $n \in \mathbb{N}^m$, $Z^{(n)}$ and Z^∞ can be immersed in the Banach space $C = (C^0([0, T], \mathbb{R}^m \times [0, 1]^{|\mathcal{S}|}), \|\cdot\|_\infty)$, where we assume that $\mathbb{R}^m \times [0, 1]^{|\mathcal{S}|}$ is equipped with the 1-norm. For each solution, $Y^{(n)}$ (resp., Y^{∞}) denotes the continuous component, evolving in $\mathbb{R}^m$, while $P^{(n)}$ (resp., P^{∞}) denotes a probability distribution over $\mathcal{S}$. Observe that for a fixed n if $x \notin \mathcal{T}_0^{(n)}$ then $P^{(n)}(x;t) = 0 \,\forall\, t \in [0,T]$.

5.1 Decomposition as a perturbed dynamical system

Before proving our result of convergence we rewrite system (15)-(16) suitably. We extend the matrix $Q^{(n)}(y)$ to a new matrix $A^{(n)}(y)$ so that (15)-(16) can be written in matrix form. To do so we define the $(m+\mathcal{N}(n)) \times (m+\mathcal{N}(n))$ block-matrix:

$$A^{(n)}(y) = \left[\begin{array}{c|c} \underline{0} & R^{(n)}(y) \\ \hline \underline{0} & Q^{(n)}(y) \end{array}\right] \tag{18}$$

where the upper-left block has dimensions $m \times m$ and $R^{(n)}(y)$ is defined component-wise as:

$$[R^{(n)}(y)]_{i,x} = \sum_{l \in \mathcal{L}(x)} \mathbb{I}_{\{x \in \partial \mathcal{T}_0^{(l,n)}\}} \mathcal{Y}_i^{(l,n)}(x) f_l(x_0 + y) \quad \text{for } i = 1, \dots, m, \text{ and } x \in \mathcal{T}_0^{(n)}.$$

Using this matrix, we can rewrite the DBP equations (15)-(16) as:

$$\frac{dZ^{(n)}}{dt} = A^{(n)}(Y^{(n)}(t))Z^{(n)}(t). \tag{19}$$

Our proof relies on the decomposition of this system into two parts: a linear one and a non-linear one acting as a perturbation. To do so, in the previous equation we rewrite $A^{(n)}(Y^{(n)})$ as $A^{(n)}(0) + \Delta A^{(n)}(Y^{(n)})$ where we have set $\Delta A^{(n)}(Y^{(n)}) = A^{(n)}(Y^{(n)}) - A^{(n)}(0)$. Then, (19) takes the form:

$$\frac{dZ^{(n)}}{dt} = (A^{(n)}(0) + \Delta A^{(n)}(Y^{(n)}(t))Z^{(n)}(t) \tag{20}$$

We will refer to this system as the *non-linear perturbed system.* Instead we define the *linear non-perturbed system* given by:

$$\frac{dZ_\Lambda^{(n)}}{dt} = A^{(n)}(0)Z_\Lambda^{(n)}(t) \tag{21}$$

whose solution we denote by $Z_\Lambda^{(n)} = [Y_\Lambda^{(n)}, P_\Lambda^{(n)}]^T \in C$.

Proof strategy. With this decomposition, we first prove convergence of the linear part in Section 5.2. Then convergence result for the full system will be proved in Section 5.3. The proofs will be based on the observation that the family of distributions $(P^{\infty}(\,\cdot\,;t) : t \in [0,T])$ is tight in $\mathcal{P}(\mathcal{S})$ [10], due to the fact that the function $P^{\infty}(\,\cdot\,;t)$ is continuous in t and we are considering a compact interval $[0,T]$, so we have that

$$\forall\, \epsilon > 0 \quad \exists\, K \subseteq \mathcal{S} \text{ compact s.t. } \sup_{t \in [0,T]} \sum_{x \notin K} P^{\infty}(x;t) \le \epsilon.$$

This property can be rephrased by considering the states outside a truncation $\mathcal{T}_0^{(n)}$ and saying that there exists a non-negative sequence $C^{(n)}$ such that:

$$\sup_{t \in [0,T]} \sum_{x \notin \mathcal{T}_0^{(n)}} P^{\infty}(x;t) \le C^{(n)} \xrightarrow{n \to \infty} 0$$

In particular we can restrict the previous sum to states on the border of a truncation (that can be viewed as states outside a suitable smaller truncation). As a result we can express the tightness property in the following form:

$$\sup_{t\in[0,T]} \sum_{x\in\partial\mathcal{T}_0^{(n)}} P^{\infty}(x;t) \le C^{(n)} \xrightarrow{n\to\infty} 0. \tag{22}$$

In the rest of this section the proofs are presented for $\mathcal{S} = \mathbb{N}^m$ since the general case $\mathcal{S} \subset \mathbb{N}^m$ can be easily derived from it. In the latter case the limit behavior as $n \longrightarrow \infty$ is intended for $n \in \mathbb{N}^m$ and $n_i \longrightarrow \sup\{x_i : x \in \mathcal{S}\}$ for all i.

5.2 Convergence of linear non-perturbed system

THEOREM 5.1. *Suppose that hypothesis (H2) holds. Then $Z_\Lambda^{(n)}$ solution of the linear non-perturbed system (21) is such that:*

$$\lim_{n\to\infty} \|Z_\Lambda^{(n)} - Z^{\infty}\|_{\infty} = 0$$

and, in particular, for some non-negative real sequence $\{\lambda^{(n)}\}_{n\in\mathbb{N}^m}$:

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|P_\Lambda^{(n)}(\,\cdot\,;t) - P^{\infty}(\,\cdot\,;t)\|_1 = 0, \tag{23}$$

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|Y_\Lambda^{(n)}(t)\|_1 \le \lim_{n\to\infty} \lambda^{(n)} = 0. \tag{24}$$

We start by proving (23). Using (8) with $y = 0$, we write explicitly the equations for $P_\Lambda^{(n)}(x,t)$:

$$\frac{dP_\Lambda^{(n)}(x)}{dt} = -\sum_{l\in\mathcal{L}} \mathbb{I}_{\{\tilde{l}^{(n)}(x)\neq 0\}} f_l(x) P_\Lambda^{(n)}(x;t) + \sum_{\substack{x'\in\mathcal{T}_0^{(n)}\\ x'\neq x}} \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x'+\tilde{l}^{(n)}(x')=x\}} f_l(x') P_\Lambda^{(n)}(x';t). \tag{25}$$

Observe that the dynamics of $P_\Lambda^{(n)}$ do not depend on $Y_\Lambda^{(n)}$ so we can solve these components independently and then plug the solution in the equations for $Y_\Lambda^{(n)}$.

We introduce a new function $\tilde{P}^{(n)}(t) = (\tilde{P}^{(n)}(x_\infty;t), \tilde{P}^{(n)}(x;t) : x \in \mathcal{T}_0^{(n)})$ whose evolution is given by the system of ODEs:

$$\frac{d\tilde{P}^{(n)}(x)}{dt} = -\sum_{l\in\mathcal{L}} f_l(x)\tilde{P}^{(n)}(x;t) + \sum_{x'\in\mathcal{T}_0^{(n)}} \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x'+l=x\}} f_l(x')\tilde{P}^{(n)}(x';t) \tag{26}$$

$$\frac{d\tilde{P}^{(n)}(x_\infty)}{dt} = \sum_{l\in\mathcal{L}} \sum_{x\in\partial\mathcal{T}_0^{(l,n)}} f_l(x)\tilde{P}^{(n)}(x;t) \tag{27}$$

with initial conditions $\tilde{P}^{(n)}(x;0) = P^{(n)}(x;0)$ and $\tilde{P}^{(n)}(x_\infty;0) = 0$.

Comparing (25) with (26)-(27) one can conclude that:

$$\tilde{P}^{(n)}(x;t) \le P_\Lambda^{(n)}(x;t) \le \tilde{P}^{(n)}(x;t) + \tilde{P}^{(n)}(x_\infty;t) \quad \forall\, t \in [0,T],\ \forall\, x \in \mathcal{T}_0^{(n)}.$$

Therefore, (23) will be proved if we prove that the following two limits hold:

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \sum_{x\in\mathbb{N}^m} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| = 0, \tag{28}$$

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \tilde{P}^{(n)}(x_\infty;t) = 0. \tag{29}$$

Let us start by proving the first limit. We have:

$$\frac{d(\tilde{P}^{(n)}(x) - P^{\infty}(x))}{dt} = -\sum_{l\in\mathcal{L}}(\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t))f_l(x) + \\ + \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x-l\in\mathcal{T}_0^{(n)}\}}(\tilde{P}^{(n)}(x-l;t) - P^{\infty}(x-l;t))f_l(x-l) + \\ + \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x-l\notin\mathcal{T}_0^{(n)}\}} P^{\infty}(x-l;t) f_l(x-l).$$

Integrating between 0 and t gives:

$$\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t) = -\int_0^t \sum_{l\in\mathcal{L}}(\tilde{P}^{(n)}(x;s) - P^{\infty}(x;s))f_l(x)\,ds + \\ + \int_0^t \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x-l\in\mathcal{T}_0^{(n)}\}}(\tilde{P}^{(n)}(x-l;s) - P^{\infty}(x-l;s))f_l(x-l)\,ds + \\ + \int_0^t \sum_{l\in\mathcal{L}} \mathbb{I}_{\{x-l\notin\mathcal{T}_0^{(n)}\}} P^{\infty}(x-l;s) f_l(x-l)\,ds.$$

We consider the absolute value and sum over $x \in \mathcal{T}_0^{(n)}$, to obtain:

$$\sum_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| \leq 2\int_0^t \sum_{l\in\mathcal{L}} \sum_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;s) - P^{\infty}(x;s)| f_l(x)\,ds + \\ + \int_0^t \sum_{l\in\mathcal{L}} \sum_{x\in\partial\mathcal{T}_0^{(l,n)}} P^{\infty}(x-l;s) f_l(x-l)\,ds.$$

Now, applying (H2) and tightness of P^{∞} (22):

$$\sum_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| \leq 2|\mathcal{L}|C \int_0^t |\tilde{P}^{(n)}(x;s) - P^{\infty}(x;s)|\,ds + |\mathcal{L}|TCC^{(n)}.$$

Applying Gronwall's inequality [10] to the scalar function $\sum_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)|$, we get:

$$\sup_{t\in[0,T]} \sum_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| \leq e^{2T|\mathcal{L}|C}|\mathcal{L}|TCC^{(n)}.$$

Finally, we can apply the tightness property of P^{∞} again to conclude that:

$$\sup_{t\in[0,T]} \sup_{x\in\mathbb{N}^m} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| \leq \sup_{t\in[0,T]} \sup_{x\in\mathcal{T}_0^{(n)}} |\tilde{P}^{(n)}(x;t) - P^{\infty}(x;t)| + \sup_{t\in[0,T]} \sum_{x\notin\mathcal{T}_0^{(n)}} P^{\infty}(x;t) \\ \leq e^{2T|\mathcal{L}|C}|\mathcal{L}|TCC^{(n)} + C^{(n)} \to 0.$$

We have thus proved that (28) holds.

Limit (29) follows easily by observing that the previous proof implies that tightness property holds also for the distribution $\tilde{P}^{(n)}$ (possibly for a different sequence $C^{(n)}$). So we have:

$$\frac{d\tilde{P}^{(n)}(x_{\infty})}{dt} = \sum_{l\in\mathcal{L}} \sum_{x\in\partial\mathcal{T}_0^{(l,n)}} \tilde{P}^{(n)}(x;t) f_l(x) \leq |\mathcal{L}|CC^{(n)} \to 0.$$

This concludes the proof of (23).

We now prove (24). In virtue of the uniform convergence of $P_\Lambda^{(n)}$ to P^∞ we can apply the tightness property also to this distribution (possibly considering a different sequence $C^{(n)}$). We have:

$$\left\| \frac{dY_\Lambda^{(n)}}{dt} \right\|_1 = \left\| \sum_{l\in\mathcal{L}} \sum_{x\in\partial\mathcal{T}_0^{(l,n)}} \mathcal{Y}^{(l,n)}(x) f_l(x) P_\Lambda^{(n)}(x;t) \right\|_1 \leq m\bar{l}|\mathcal{L}|CC^{(n)} \to 0, \text{ where } \bar{l} = \max_{l\in\mathcal{L}} \max_{i=1,\ldots,m} |l_i|.$$

This concludes the proof of Theorem 5.1.

5.3 Convergence result

We now proceed to prove the convergence of the non-linear perturbed system by exploiting the results already proved for the linear counter-part. Our main result is stated in the following theorem.

THEOREM 5.2. *Suppose that hypotheses (H1) and (H2) hold. Then $Z^{(n)}$ solution of the non-linear perturbed system (20) satisfies:*

$$\lim_{n\to\infty} \|Z^{(n)} - Z^\infty\|_\infty = 0$$

and, in particular

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|P^{(n)}(\,\cdot\,;t) - P^\infty(\,\cdot\,;t)\|_1 = 0, \tag{30}$$

$$\lim_{n\to\infty} \sup_{t\in[0,T]} \|Y^{(n)}(t)\|_1 = 0. \tag{31}$$

We start by proving the following proposition.

PROPOSITION 5.3. *Hypothesis (H1) implies that there exists a constant $K > 0$ such that:*

$$\|\Delta A^{(n)}(y)\|_1 \leq K\|y\|_1 \quad \forall\, y \in \mathbb{R}^m_{\geq 0}, \forall\, n \in \mathbb{N}^m$$

PROOF. Using (18) we can estimate the 1-norm of $\Delta A^{(n)}(y)$ as:

$$\|\Delta A^{(y)}(y)\|_1 = \max_{j\in\{1,\ldots,m\}\cup\mathcal{T}_0^{(n)}} \sum_{i\in\{1,\ldots,m\}\cup\mathcal{T}_0^{(n)}} |[\Delta A^{(n)}((y))]_{i,j}| \leq (m\bar{v}+2) \max_{x\in\mathcal{T}_0^{(n)}} \sum_{l\in\mathcal{L}} |f_l(y+x) - f_l(x)|$$

where the last inequality follows from the fact that, for a fixed column indexed by $x \in \mathcal{T}_0^{(n)}$, we have that the term $m\bar{v}\sum_{l\in\mathcal{L}} |f_l(y+x) - f_l(x)|$ takes into account the components of the submatrix $R^{(n)}(y)$ and the term $2\sum_{l\in\mathcal{L}} |f_l(y+x) - f_l(x)|$ takes into accounts the components in the submatrix $Q^{(n)}(y)$ (observe that at most $|\mathcal{L}|$ of the non-diagonal entries of a given column are non-zero, since they correspond to the reactions starting in the observed state indexing that column).

Applying (H1) to the r.h.s. of the previous inequality our assertion follows immediately. □

It is possible to verify by direct derivation that $Z^{(n)}(t)$ can be expressed as:

$$Z^{(n)}(t) = Z_\Lambda^{(n)}(t) + \int_0^t e^{(t-s)A^{(n)}(0)} \Delta A^{(n)}(Y^{(n)}(s)) Z^{(n)}(s)\, ds. \tag{32}$$

The operator $e^{tA^{(n)}(0)}$ is uniformly bounded under our current hypotheses , i.e., for a constant $M > 0$ we have

$$\|e^{tA^{(n)}(0)}\| \leq M \quad \forall\, t \in [0,T], \forall n \in \mathbb{N}^m.$$

This can be proved using an argument similar to that used in the proof of Proposition 5.3, so that we have:

$$\|A^{(n)}(0)\|_1 = \max_{j\in\{1,\ldots,m\}\cup\mathcal{T}_0^{(n)}} \sum_{i\in\{1,\ldots,m\}\cup\mathcal{T}_0^{(n)}} |[A^{(n)}(0)]_{i,j}| \leq m\bar{l}|\mathcal{L}|C + |\mathcal{L}|C + |\mathcal{L}|C \leq (m\bar{l}+2)|\mathcal{L}|C.$$

Let us now prove (31). Let us fix a constant $H \gg 2$ and define:

$$\Delta_1^{(n)} = \sup\left\{\delta \geq 0 : \sup_{s\in[0,T]} \|Z^{(n)}(s)\|_1 \leq H\right\}.$$

For n sufficiently large, we have that $\|Z^{(n)}(0)\|_1 = 1$ so $\Delta_1^{(n)} > 0$. Suppose $\Delta_1^{(n)} < +\infty$, then, by continuity, $\|Z^{(n)}(\Delta_1^{(n)})\|_1 = H$. Let n sufficiently large be fixed and consider $0 \leq t \leq \Delta_1^{(n)}$. From (32), we have:

$$\|Y^{(n)}(t)\|_1 \leq \|Y_\Lambda^{(n)}(t)\|_1 + tMKH \sup_{s\in[0,t]} \|Y^{(n)}(s)\|_1.$$

Now, suppose that $\liminf_n \Delta_1^{(n)} \in \left[0, \frac{1}{4MKH}\right]$, then for a suitable subsequence n_k we have $\Delta_1^{(n_k)} < \frac{1}{2MKH} \quad \forall\, k$. Then

$$\sup_{t\in[0,\Delta_1^{(n_k)}]} \|Y^{(n_k)}(t)\|_1 \leq \lambda^{(n_k)} + \frac{1}{2} \sup_{s\in[0,\Delta_1^{(n_k)}]} \|Y^{(n)}(s)\|_1$$

and

$$\sup_{t\in[0,\Delta_1^{(n_k)}]} \|Y^{(n_k)}(t)\|_1 \leq 2\lambda^{(n_k)} \xrightarrow{n\to\infty} 0$$

which contradicts the continuity hypothesis according to which $\|Z^{(n)}(\Delta_1^{(n)})\|_1 = H \gg 2$, since this implies $\|Y^{(n)}(\Delta_1^{(n)})\|_1 > H - 1 \gg 1$. We have thus proved that $\lim_n \inf \Delta_1^{(n)} > \frac{1}{4MKH}$.

We can then set $t_1^* = \frac{1}{4MKH}$, for which we have, provided n is sufficiently large:

$$\sup_{t\in[0,t_1^*]} \|Y^{(n)}(t)\|_1 \leq \lambda^{(n)} + \frac{1}{4} \sup_{t\in[0,t_1^*]} \|Y^{(n)}(s)\|_1$$

and

$$\sup_{t\in[0,t_1^*]} \|Y^{(n)}(t)\|_1 \leq 2\lambda^{(n)} \xrightarrow{n\to\infty} 0.$$

Now, if $t_1^* \geq T$ we have completed our proof, if not we need to iterate the argument to extend the interval. To do this we define:

$$\Delta_2^{(n)} = \sup\left\{\delta \geq 0 : \sup_{s\in[t_1^*,t_1^*+\delta]} \|Z^{(n)}(s)\|_1 \leq H\right\}.$$

Observe that because of the previous step, for n sufficiently large we have $\|Z^{(n)}(t_1^*)\|_1 \leq 2\lambda^{(n)} \ll H$ so, for such values of n, $\Delta_2^{(n)} > 0$. Moreover, by continuity, if $\Delta_2^{(n)} < \infty$, $\|Z^{(n)}(t_1^* + \Delta_2^{(n)})\|_1 = H$. Now for $t_1^* \leq t \leq t_1^* + \Delta_2^{(n)}$ and for n sufficiently large:

$$\begin{aligned}\|Y^{(n)}(t)\|_1 &\leq \|Y_\Lambda^{(n)}(t)\|_1 + t_1^* MKH \sup_{s\in[0,t_1^*]} \|Y^{(n)}(s)\|_1 + (t - t_1^*)MKH \sup_{s\in[t_1^*,t]} \|Y^{(n)}(s)\|_1 \\ &\leq \|Y_\Lambda^{(n)}(t)\|_1 + \frac{1}{2} 2\lambda^{(n)} + (t - t_1^*)MKH \sup_{s\in[t_1^*,t]} \|Y^{(n)}(s)\|_1.\end{aligned}$$

Using an argument similar to the one used before we can prove that $\liminf_n \Delta_2^{(n)} > \frac{1}{4MKH}$, and set $t_2^* = t_1^* + \frac{1}{4MKH} = 2t_1^*$ so that

$$\sup_{t\in[0,t_2^*]} \|Y^{(n)}(t)\|_1 \leq 4\lambda^{(n)}.$$

If $t_2^* \geq T$ we can conclude, if not, we can reiterate the procedure as many times as needed. Since each time the interval is extended by a constant quantity, the whole $[0, T]$ interval is covered by a finite number h of iterations, and we get:

$$\sup_{t \in [0,T]} \|Y^{(n)}(t)\|_1 \leq 2^h \lambda^{(n)} \longrightarrow 0.$$

Let us now prove (30). This follows easily by observing that from (32), and from and the convergence of $Y^{(n)}$, we have

$$\begin{aligned} \sup_{t \in [0,T]} \|P^{(n)}(\,\cdot\,; t) - P_\Lambda^{(n)}(\,\cdot\,; t)\|_1 &\leq TKM \sup_{t \in [0,T]} \|Y^{(n)}(t)\|_1 \left(1 + \sup_{t \in [0,T]} \|Y^{(n)}(t)\|_1\right) \\ &\leq TMK(2^h \lambda^{(n)})(2^h \lambda^{(n)} + 1) \xrightarrow{n \to \infty} 0. \end{aligned}$$

Since by the previous theorem we have already proved the uniform convergence of $P_\Lambda^{(n)}$ to $P^{(n)}$, this concludes the proof of Theorem 5.2.

6 EXAMPLES

In this section we apply DBP to three examples, which are density-dependent Markov population processes with a scaling parameter N. In these cases, as discussed, the mean-field solution $x(t)$ represents the limit behavior of the re-scaled population process $X^N(t)/N$. Thus, for any given N, we will compare the DBP refinement of the unscaled process $\mathbb{E}[X^N(t)]$ against the mean-field approximation of expected populations given by $Nx(t)$. As ground-truth we take the average trajectory using Gillespie's direct stochastic simulation algorithm [16]. The stopping criterion was chosen as follows: along the simulated time horizon we fix 100 equidistant time steps, and at each time step we compute the 95% confidence interval. If semi-amplitude of the interval is greater than 1% of the mean at that step the number of simulations is increased by 1k runs or 10k runs (if the number of simulations has already exceeded 50k). This procedure is applied to all observed outputs.

Stochastic simulation as well as the numerical solutions of the mean-field equations and the DBP equations were performed in Matlab. The DBP equations were generated from a prototype implementation within the (Java-based) software tool ERODE [5]. In the runtime comparisons we do not report the time taken to generate the DBP Matlab file because we found it to be negligible with respect to the solution time. The expansions from [12] where computed using the Python-based implementation therein reported. Since the considered implemented computes both expansions simultaneously. All experiments were conducted on a laptop equipped with a 2.8 GHz Intel i7 quad-core processor and 16 GB RAM.

6.1 Coxian Queuing Systems

Two-phase Coxian distribution. We first consider the example anticipated in Section 1, an $M/Cox/N$ queuing system with Poisson arrivals with rate $N\lambda$ and service rate with a two-phase Coxian distribution. The three degrees of freedom of the Coxian distribution are identified by parameters p (probability of transitioning from first to second phase of service), and μ_1, μ_2 (exponential rates at each stage). Following [2], a service-time distribution with mean E and variance V can be obtained by setting $p = E^2/(2V)$, $\mu_1 = 2/E$ and $\mu_2 = E/V$. Using a standard state space description [2], the queuing system can be represented as a Markov population process with state $x = (x_{Q_1}, x_{Q_2}, x_S)$ where: x_{Q_1} is the number of jobs requiring the first phase of service; x_{Q_2} is the number of jobs in second phase; and x_S is the number of servers available for the first phase of service. Thus, the queue length in state x is $x_{Q_1} + x_{Q_2}$. For a system with N independent servers, the jump vectors

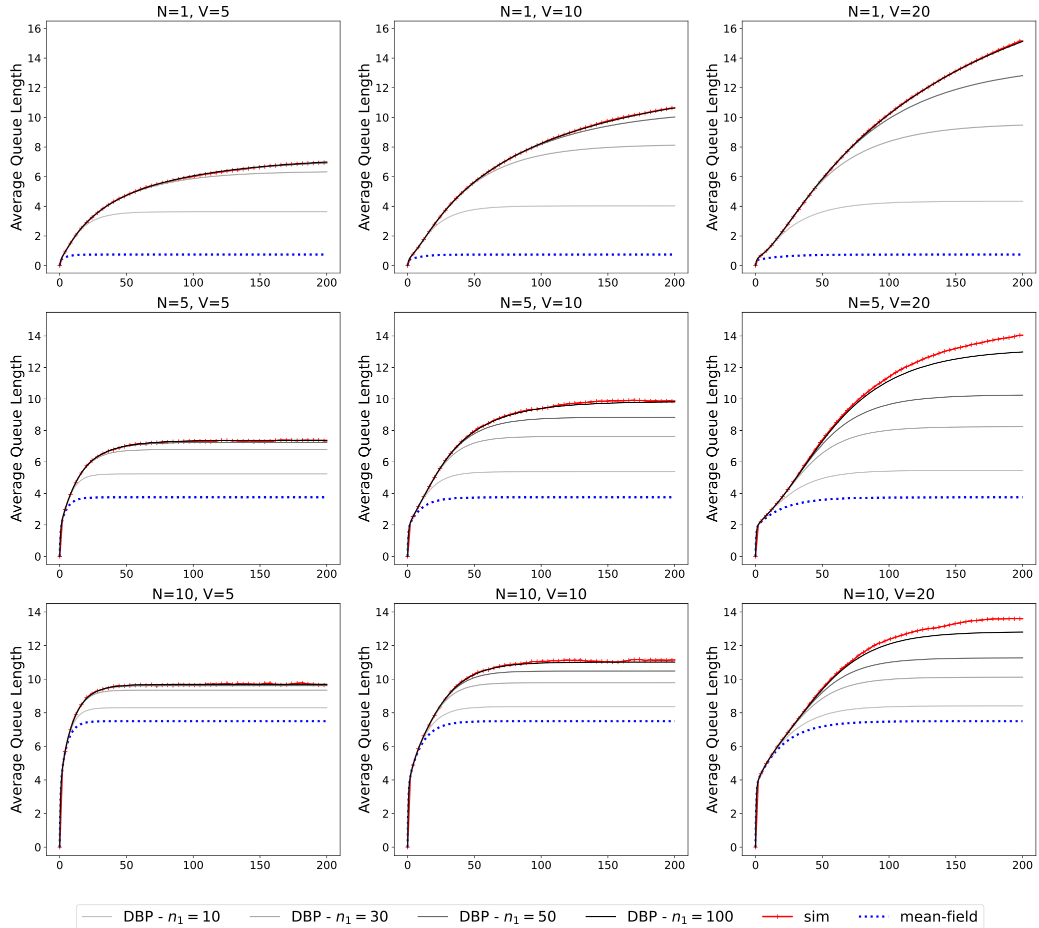

Fig. 2. DBP applied to a $M/Cox/N$ queuing system where the truncation is applied to n_1 jobs waiting for service in the first phase of a two-phase Coxian distribution.

and the associated transition functions are given by:

$$\begin{aligned}
l_1 &= +e_{Q_1}, && \text{at rate } N\lambda, \\
l_2 &= -e_{Q_1} - e_S + e_{Q_2} && \text{at rate } p\mu_1 \min(x_{Q_1}, x_S), \\
l_3 &= -e_{Q_1} && \text{at rate } (1-p)\mu_1 \min(x_{Q_1}, x_S), \\
l_4 &= -e_{Q_2} + e_S && \text{at rate } \mu_2(N - x_S),
\end{aligned}$$

where e_i denotes the canonical vector which has the i-th coordinate equal to one. We set the initial condition $\bar{x}_0 = (0, 0, N)$. The mean-field approximation has a Lipschitz continuous but non-differentiable drift due to the presence of the mininum in the transition functions, for which available results of mean-field refinement in [12] are not applicable.

Figure 2 compares the transient evolution of average queue lengths for $\lambda = 0.75$ and different variances (using a unitary mean service time as in Table 1) and for various truncations in the form $n = (n_1, N, N)$, i.e., by truncating the number of jobs requiring first-phase service to n_1. In general,

Table 2. Runtimes (in s) for simulations (SIM) and DBP with bound (n_1, N, N) for the two-phase Coxian distribution; the number of equations refer to the size of the resulting DBP models. The solution of the mean-field system takes 0.06s on average.

N	V	*SIM*		*DBP*			
				$n_1 = 10$	$n_1 = 30$	$n_1 = 50$	$n_1 = 100$
		Time	*# runs*	*Time*	*Time*	*Time*	*Time*
				(22 eqs.)	(62 eqs.)	(102 eqs.)	(202 eqs.)
1	5	204	130k	0.07	0.08	0.14	0.19
1	10	224	140k	0.08	0.10	0.14	0.21
1	20	213	140k	0.12	0.22	0.27	0.40
				(66 eqs.)	(186 eqs.)	(306 eqs.)	(606 eqs.)
5	5	574	70k	0.37	0.85	1.32	2.58
5	10	944	100k	0.45	0.90	1.43	2.71
5	20	1331	130k	0.53	0.97	1.56	2.82
				(121 eqs.)	(341 eqs.)	(561 eqs.)	(1111 eqs.)
10	5	487	29k	0.93	2.72	4.37	8.85
10	10	889	60k	0.93	2.90	4.54	9.22
10	20	1622	100k	1.02	3.01	4.53	9.24

the mean-field approximation does not behave satisfactorily and, as mentioned in Section 1, provides estimates for long enough time horizons that are insensitive to the service time distribution. DBP can refine the mean in all cases, although larger values of n_1 are needed to improve the accuracy with larger variances of the service-time distribution.

Table 2 reports the runtimes as well as the size of the resulting system of equations when applying DBP. As expected, the computational cost of the DBP analysis grows with n_1 and is larger than that of the mean-field solution. However it can refine the mean estimate at a small fraction (no more than about 2% across all cases) of the time taken to perform stochastic simulation.

Larger number of phases. To show how our method behaves with larger models, we also present examples of $M/Cox/N$ queues using a Coxian distribution with more phases. A K-phase Coxian distribution is defined by the parameter vectors $(p_1, \ldots, p_{K-1})$ and $(\mu_1, \ldots, \mu_K)$. Thus, the queuing system is represented by a Markov population process with state descriptor $x = (x_{Q_1}, \ldots, x_{Q_K}, x_{S_1}, \ldots x_{S_{K-1}})$. For N independent servers we define the following transition functions:

$$
\begin{aligned}
a_1 &= +e_{Q_1}, && \text{at rate } N\lambda,\\
d_1 &= -e_{Q_1} && \text{at rate } (1-p_1)\mu_1 \min(x_{Q_1}, x_{S_1}),\\
&\ldots\\
a_i &= -e_{Q_{i-1}} - e_{S_{i-1}} + e_{Q_i} + e_{S_i} && \text{at rate } p_{i-1}\mu_{i-1} \min(x_{Q_{i-1}}, x_{S_{i-1}}),\\
d_i &= -e_{Q_i} - e_{S_i} + e_{S_1} && \text{at rate } p_i\mu_i \min(x_{Q_i}, x_{S_i}),\\
&\ldots\\
a_K &= -e_{Q_{K-1}} - e_{S_{K-1}} + e_{Q_K} && \text{at rate } p_{K-1}\mu_{K-1} \min(x_{Q_{K-1}}, x_{S_{K-1}})\\
d_K &= -e_{Q_K} + e_{S_1} && \text{at rate } \mu_K \min(x_{Q_K}, N - x_{S_1} - \ldots - x_{S_{K-1}}).
\end{aligned}
$$

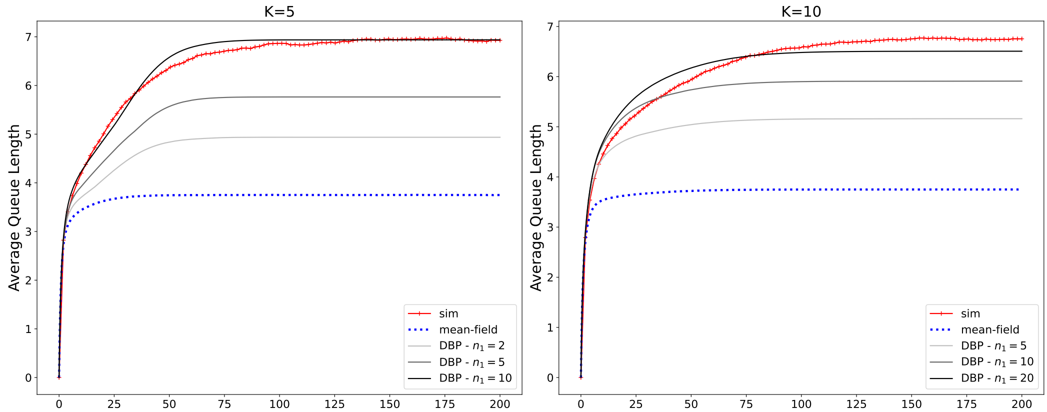

Fig. 3. DBP applied to the $M/Cox/N$ queuing systems with $K = 5$ and $K = 10$ phases for the Coxian service-time distribution.

Table 3. Runtimes (in s) for simulations (SIM), mean-field (MF) and DBP for the $M/Cox/N$ queuing system with $N = 5$ and Coxian distributed service times with $K = 5$ and $K = 10$ phases matching unitary mean and variance $V = 5$.

K	*Method*	*Parameters*	*Time*
	SIM	80k runs	604
	MF	-	0.22
5	*DBP*	$n_1 =$ 2, 57 eqs.	0.41
	DBP	$n_1 =$ 5, 105 eqs.	1.04
	DBP	$n_1 =$ 10, 185 eqs.	1.95
	SIM	100k runs	953
	MF	-	0.36
10	*DBP*	$n_1 =$ 5, 67 eqs.	0.43
	DBP	$n_1 =$ 10, 107 eqs.	0.79
	DBP	$n_1 =$ 20, 187 eqs.	2.20

Figure 3 shows the results for $K = 5$ and $K = 10$ with $N = 5$ and Coxian parameters manually tuned so to have unitary mean and variance $V = 5$ (see Appendix C). For $K = 5$ we used bounds in the form $n = (n_1, 1, 1, 1, 1, 0, 0, 0, 0)$, with $n_1 = 2, 5, 10$; for $K = 10$ we used the bounds $n_1 = 5, 10, 20$, $n_2 = \ldots = n_4 = 1, n_5 = n_{19} = 0$. Runtimes are reported in Table 3. Since we are using modest truncations, even if the computational time is larger than that of the classic mean-field approximation, for each of the chosen bounds it does not exceed 2.2s. Nonetheless, in Figure 3 it is possible to see that the approximation is consistently improved.

6.2 Malware Propagation Model

Here we consider the malware propagation model from [1, 12, 18]. It is composed of N nodes, where each node can be dormant (D), active (A) or susceptible (S). Since the total number of nodes is constant, the model can be described by the state vector $x = (x_D, x_A)$ where the number of susceptible nodes at each state is given by $N - x_D - x_A$. The process evolves according to the

Table 4. Runtimes (in s) for simulations (SIM), mean-field (MP), DBP and expansions in $1/N$ and $1/N^2$ (EXP) for the Malware Propagation Model with $N = 50$ agents.

Model	δ	*SIM*		*MF*	*DBP* $n = (24, 24)$		*EXP* $1/N$	$1/N^2$
		Time	*#runs*	*Time*	*Time*	*#eqs.*	*Time*	*Time*
Stable	0.5	100	21k	0.06	0.83	628	0.07	0.42
Unstable	0.1	2819	380k	0.08	2.17	628	0.06	0.39

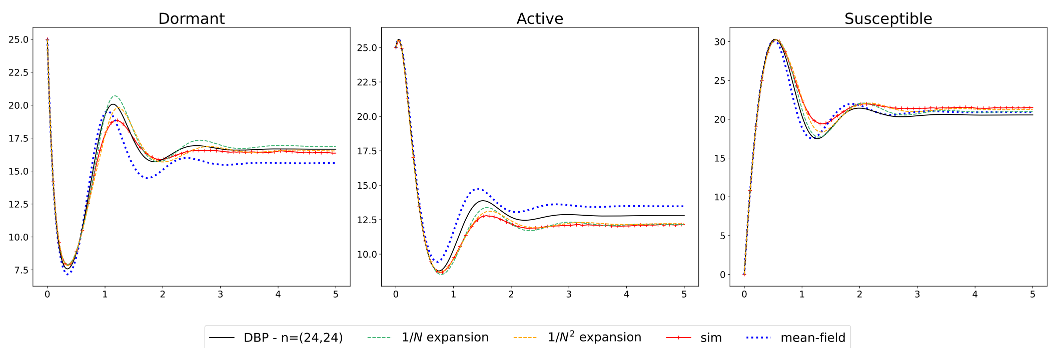

Fig. 4. Numerical results for the malware propagation model with unique attractor orbit ($\delta = 0.50$) for mean-field approximation, $1/N$ and $1/N^2$ expansions in [12], and DBP.

following transitions and jump vectors:

$$
\begin{aligned}
l_1 &= -e_D + e_A && \text{at rate } \left(1 + \frac{10x_A}{x_D + N\delta} x_D\right), \\
l_2 &= -e_A && \text{at rate } 5x_A, \\
l_3 &= +e_D && \text{at rate } \left(\beta + \frac{10}{N} x_D\right)(N - x_D - x_A).
\end{aligned}
$$

In [12] it is discussed that there exists a parameter value $\delta^* \approx 0.18$ such that for $\delta > \delta^*$ the mean field approximation has a unique attractor; instead, for $\delta < \delta^*$ the ODE has an orbit cycle, which may cause significant approximation errors.

Stable model. To show the error behavior in the case of a stable mean-field approximation, we set $N = 50, \delta = 0.5, \beta = 0.1$, and $\bar{x}_0 = (\bar{x}_D, \bar{x}_A) = (25, 25)$, as done in [12]. Figure 4 compares DBP applied with bound $n = (24, 24)$, against the $1/N$ and $1/N^2$ expansion approximations in [12]. We found that these methods all perform comparably with each other in terms of accuracy, improving the mean-field estimation (especially for the average population of dormant nodes), while the $1/N^2$ expansion is faster than DBP (see Table 4).

Unstable model. Using $\delta = 0.1$, Figure 5 shows the presence of an orbit cycle in the mean-field approximation that causes instability to both the expansions $1/N$ and $1/N^2$. Instead, using the same n as in the previous case, DBP exhibits damped oscillatory behaviour that refines the mean-field approximation, especially over longer time horizons when the oscillations tend to fade away. Runtimes can be found in Table 4.

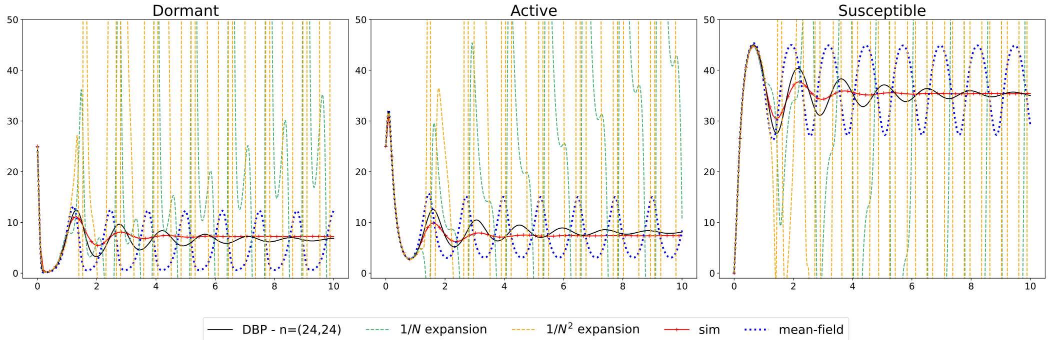

Fig. 5. Numerical results for the malware propagation model with orbit cycle ($\delta = 0.10$) for mean-field approximation, $1/N$ and $1/N^2$ expansions in [12], and DBP.

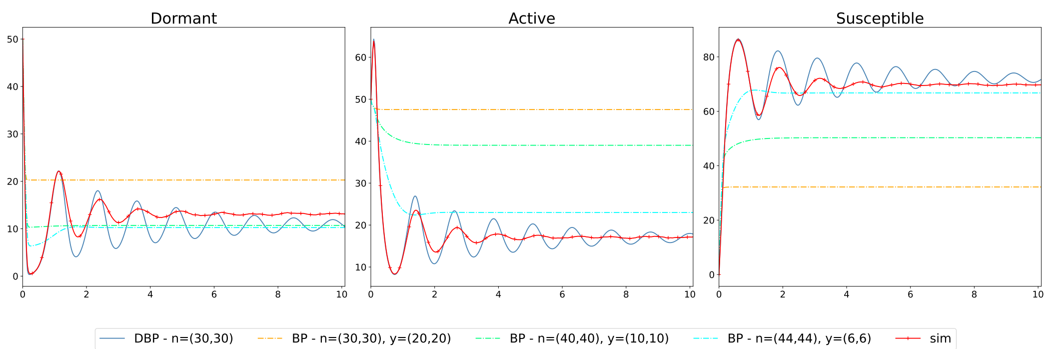

Fig. 6. Numerical results for the unstable malware propagation model ($\delta = 0.1$) and DBP with bound $n = (15, 15)$ and BP on different truncations $\mathcal{T}_y^{(n)}$.

Impact of dynamic shift of truncation. The structure of this model allows us to show the impact of the dynamic shift. In order to do so, we consider the unstable model with $N = 100$, $\delta = 0.1$, $\beta = 0.2$ and compare DBP against different BPs that are not modulated by the mean-field approximation, i.e., solutions to systems in the form of (9). Figure 6 shows the results of DBP with $n = (30, 30)$ and BP applied on a truncated state space $\mathcal{T}_y^n$ with different values of n and y (defined as in (3)). We use initial condition $\bar{x}_0 = (\bar{x}_D, \bar{x}_A) = (50, 50)$. Observe that with this choice, when applying BP we are forced to choose a truncation containing the initial state.

The results indicate that, for the same size of the truncated state space $n = (30, 30)$, DBP significantly outperforms BP. Since the master equations for the truncated state space have the same functional form in DBP and BP, the difference in behavior is due to the dynamic shift of the truncation by the coupling with the two mean-field equations that approximate the average population of dormant and active nodes. To achieve comparable accuracy with DBP, BP requires much larger state spaces—i.e., using $n = (44, 44)$, corresponding to over a twofold increase of the truncated state space size (2025 states for BP against 961 states for DBP).

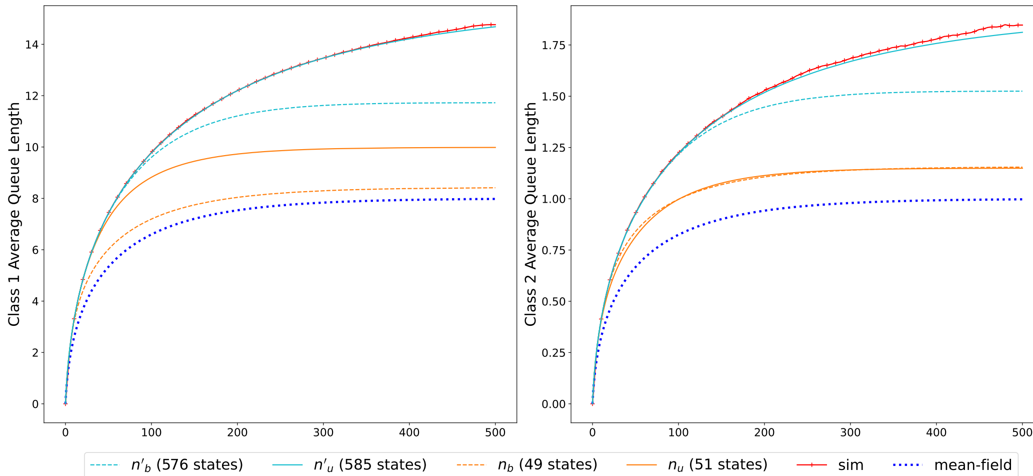

Fig. 7. Numerical results for the egalitarian processor sharing queuing system with $K = 2$ classes using different DBP truncation parameters.

6.3 Egalitarian Processor Sharing

We now discuss a simplified version of a queuing system with generalized processor sharing proposed in [29, 30]. We apply an *egalitarian* policy whereby all customers are assigned the same weight; in particular we use the rate functions adopted in [43]. A system with K classes of customers can be represented as a Markov population processes with state $x = (x_{Q_1}, \dots, x_{Q_K})$, where each component identifies a class of customers in the queue. The jump vectors and the transition functions associated with the process are:

$$a_i = +e_{Q_i}, \qquad \text{at rate } N\lambda_i,$$

$$d_i = -e_{Q_i}, \qquad \text{at rate } \mu\frac{x_{Q_i}}{\sum_{j=1}^{K} x_{Q_j} + N},$$

for $i = 1, \dots, K$.

We consider models with $K = 2, 3, 4$ classes, always setting the initial condition to $\bar{x}_0 = (0, \dots, 0)$ and $N = 1$. For $K = 2$ classes, we first show how the choice of the bound n can impact the accuracy of the approximation. In fact, when the arrival rates λ_i are different we expect that the probability mass in each truncation will be concentrated on states with $x_{Q_i} > x_{Q_j}$ for $\lambda_i > \lambda_j$; specifically, we expect that the ratio between the mean number of class-i customers and the mean number of class-j customers is approximately equal to $\frac{\lambda_i}{\lambda_j}$. This suggests a heuristic of *unbalanced* bounds where each dimension is set proportionally to the arrival rate of each class. We will show that applying this heuristic, instead of considering balanced bounds of the form $n_i = n_j$ for any i and j, improves the approximation. In particular we will show a direct comparison for $K = 2$ and then apply this heuristic also for $K = 3$ and $K = 4$.

Case $K = 2$. To show the impact of the choice of n, we take two pairs of vectors. The pair $n_b = (6, 6)$ and $n'_b = (23, 23)$ considers two truncations of increasing size which are balanced, i.e., $n_{Q_1} = n_{Q_2}$; the pair $n_u = (16, 2)$ and $n'_u = (64, 8)$ considers two truncations where $n_{Q_1} = 8n_{Q_2}$. These values have been chosen such that the number of states in the truncated state spaces for n_b and n_u,

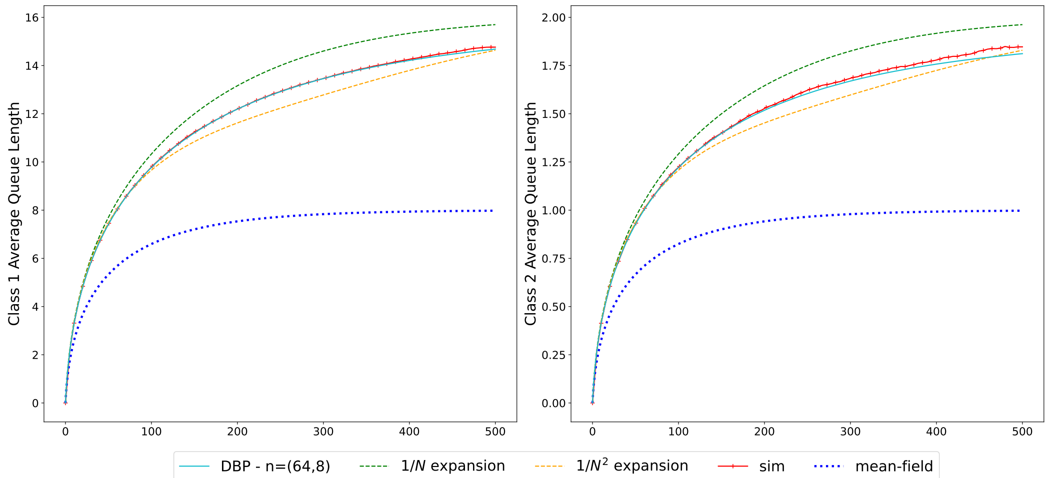

Fig. 8. Comparison between DBP with $n = (32, 4)$ and the $1/N$, $1/N^2$ expansions from [12] for the queuing system with egalitarian process sharing presented in Section 6.3.

as well as those for n'_b and n'_u, are similar. This way, the difference in the accuracy may be more directly related to the choice of the truncation. Figure 7 shows the results of the approximations for the average queue lengths for both class of customers. They indicate that using a balanced truncation leads to less accurate approximations than an unbalanced truncation with comparable number of states. Nevertheless, every DBP approximation improves the mean-field estimates in all cases considered here.

Using N as scaling parameter, this model admits a mean-field limit with a differentiable drift; therefore it is possible to apply the $1/N$ and $1/N^2$ refinements proposed in [12]. Figure 8 compares the refinements against the most accurate DBP approximation shown in Fig. 7, i.e., $n'_u = (64, 8)$. With this choice of bound, DBP is superior to both refinements across the transient regime. As the trajectories approach stationarity, the accuracy of the $1/N^2$ refinement and DBP become comparable. The runtime comparison, reported in Table 5, shows that the $1/N^2$ expansion requires less computational time (for $K = 2$ as well as in the forthcoming cases).

Case $K = 3, 4$. For $K = 3$ we chose $\lambda_1 = 0.8, \lambda_2 = 0.4, \lambda_3 = 0.1, \mu = 1.5$. Following the same argument applied for $K = 2$, we chose the bounds for Q_1, Q_2, Q_3 proportionally, so we set $n = (16, 8, 2), (24, 12, 4), (32, 16, 4)$, and $(36, 18, 5)$. For $K = 4$ and rates constants $\lambda_1 = 0.8, \lambda_2 = 0.4, \lambda_3 = 0.2, \lambda_4 = 0.1, \mu = 2$ we chose $n = (8, 4, 2, 1), (12, 6, 3, 2), (16, 8, 4, 2)$, and $(20, 10, 5, 3)$. Results for the average queue length are shown in Figure 9 while the results for per-class queue lengths are reported in Appendix B. In these cases DBP performs comparably to the $1/N^2$ expansion, although it tends to better approximate the transient evolution. Regarding runtimes, the best DBP approximations take significantly longer than the $1/N^2$ expansions; however DBP is still very competitive with respect to stochastic simulation, always requiring less than 3% runtime.

Table 5. Runtimes (in s) for simulations (SIM), mean-field (MF), DBP and $1/N^2$ expansion (EXP) for the egalitarian processor sharing queuing system with K = 2,3, and 4 classes of customers.

K	*Method*	*Parameters*	*Time*
2	*SIM*	150k runs	924
	MF	-	0.07
	DBP	$n = (6, 6)$, 54 eqs.	0.10
	DBP	$n = (16, 2)$, 56 eqs.	0.08
	DBP	$n = (23, 23)$, 581 eqs.	0.69
	DBP	$n = (64, 8)$, 590 eqs.	0.74
	EXP	$1/N$	0.05
	EXP	$1/N^2$	0.32
3	*SIM*	180k runs	1209
	MF	-	0.07
	DBP	$n = (16, 8, 2)$, 466 eqs.	0.80
	DBP	$n = (24, 12, 3)$, 1307 eqs.	13.65
	DBP	$n = (32, 16, 4)$, 2812 eqs.	27.25
	DBP	$n = (36, 18, 5)$, 4225 eqs.	31.15
	EXP	$1/N$	0.25
	EXP	$1/N^2$	3.88
4	*SIM*	360k runs	1192
	MF	-	0.04
	DBP	$n = (8, 4, 2, 1)$, 279 eqs.	0.40
	DBP	$n = (12, 6, 3, 2)$, 1101 eqs.	7.34
	DBP	$n = (16, 8, 4, 2)$, 2304 eqs.	14.19
	DBP	$n = (20, 10, 5, 3)$, 5553 eqs.	31.37
	EXP	$1/N$	0.17
	EXP	$1/N^2$	4.89

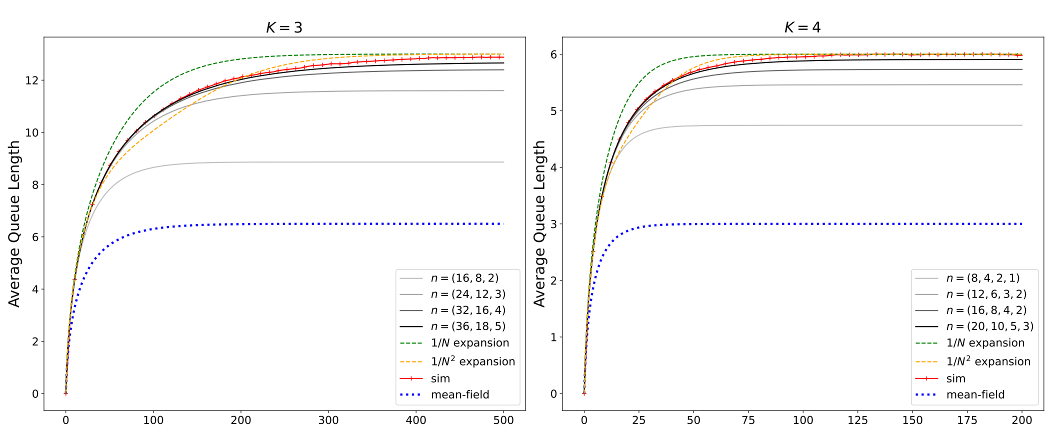

Fig. 9. Numerical results for estimation of the average queue length of the egalitarian processor sharing queuing system with $K = 3$ and $K = 4$ classes of customers.

7 CONCLUSIONS AND FUTURE WORKS

We have presented dynamic boundary projection (DBP), a new method to refine mean-field approximations of Markov population processes. It is based on a system of differential equations for the transient probability distribution of a truncation of the state space which is modulated by a mean-field approximation that essentially shifts such truncation across the whole state space. DBP is parameterized by a vector n that defines the size of the state space truncation. Thus, it leads to a family of approximations indexed by n for a given Markov population process. Importantly, for each distinct n one needs to solve a different system of DBP equations in general. In this respect it is different from related work on this subject. In particular, the method by Gast *et al.* computes the constants associated with the terms $1/N$ and $1/N^2$ of expansions of the mean-field equation for density-dependent Markov processes [12], where N is the scaling parameter. These constants correct the mean-field approximations for all N. Another difference with respect to the related work is that DBP neither makes scaling assumptions nor requires differentiability of the drift of the mean-field approximation.

The advantage of DBP is that the freedom in choosing n can be exploited to improve the accuracy of the approximations. This has been used in the analysis of the malware propagation model [1, 12, 18] where DBP has been shown to avoid instabilities exhibited with $1/N$ and $1/N^2$ size expansions. The main disadvantage is that, since DBP relies on truncations of the state space, it is still subject to the well-known *curse of dimensionality* that affects Markov population processes. For systems with many dimensions, the values of the parameter n must be kept small to avoid large computational times. However, the numerical results have shown that, even for modest truncations, the classical mean-field estimations are improved. Suitable heuristics for the choice of n can help mitigate this problem, as we discussed in the example of the queuing system with egalitarian processor sharing. However, these heuristics are model-dependent and in general it is not clear how to fix n to achieve better approximations. We plan on investigating this problem in the future.

The main theoretical result that we report here states asymptotic convergence to the original population process when n tends to infinity (uniformly over a finite time horizon). As with many convergence results, however, this does not give directly useful insights as to the behavior of the approximation for finite n. Hence natural questions that arise may be based on establishing analogies with the literature on mean-field refinements, in particular to study the following: i) whether DBP can be used to compute mean estimates of an arbitrary functional of a Markov population process; ii) whether suitable conditions hold to extend convergence to the steady state; iii) whether it is possible to establish rates of convergence or error bounds. The answers to such questions appear unclear currently, and are left as future work.

ACKNOWLEDGMENTS

This work was partially supported by the Italian Ministry for Research through the PRIN project SEDUCE, no. 2017TWRCNB.

REFERENCES

[1] Michel Benaim and Jean-Yves Le Boudec. 2008. A class of mean field interaction models for computer and communication systems. *Performance evaluation* 65, 11-12 (2008), 823–838.

[2] Gunter Bolch, Stefan Greiner, Hermann de Meer, and Kishor Trivedi. 2005. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications.* Wiley.

[3] Luca Bortolussi, Jane Hillston, Diego Latella, and Mieke Massink. 2013. Continuous approximation of collective system behaviour: A tutorial. *Performance Evaluation* 70, 5 (2013), 317–349.

[4] Maury Bramson, Yi Lu, and Balaji Prabhakar. 2010. Randomized Load Balancing with General Service Time Distributions. In *Proceedings of the ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems.* 275–286.

[5] Luca Cardelli, Mirco Tribastone, Max Tschaikowski, and Andrea Vandin. 2017. ERODE: A tool for the evaluation and reduction of ordinary differential equations. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*. Springer, 310–328.

[6] Fabio Cecchi, Sem C Borst, and JSH van Leeuwaardena. 2015. Mean-field analysis of ultra-dense csma networks. *ACM SIGMETRICS Performance Evaluation Review* 43, 2 (2015), 13–15.

[7] Augustin Chaintreau, Jean-Yves Le Boudec, and Nikodin Ristanovic. 2009. The age of gossip: spatial mean field regime. 37, 1 (2009), 109–120.

[8] R.W.R. Darling and J.R. Norris. 2008. Differential equation approximations for Markov chains. *Probab. Surveys* 5 (2008), 37–79. https://doi.org/10.1214/07-PS121

[9] Khanh N Dinh and Roger B Sidje. 2016. Understanding the finite state projection and related methods for solving the chemical master equation. *Physical biology* 13, 3 (2016), 035003.

[10] Stewart N Ethier and Thomas G Kurtz. 2009. *Markov processes: characterization and convergence*. Vol. 282. John Wiley & Sons.

[11] Nicolas Gast. 2017. Expected values estimated via mean-field approximation are $1/N$-accurate. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 1, 1 (2017), 1–26.

[12] Nicolas Gast, Luca Bortolussi, and Mirco Tribastone. 2019. Size expansions of mean field approximation: Transient and steady-state analysis. *Performance Evaluation* 129 (2019), 60–80.

[13] Nicolas Gast and Gaujal Bruno. 2010. A mean field model of work stealing in large-scale systems. *ACM SIGMETRICS Performance Evaluation Review* 38, 1 (2010), 13–24.

[14] Nicolas Gast and Benny Van Houdt. 2015. Transient and steady-state regime of a family of list-based cache replacement algorithms. In *Proceedings of the 2015 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*. 123–136.

[15] Nicolas Gast and Benny Van Houdt. 2017. A refined mean field approximation. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 1, 2 (2017), 1–28.

[16] Daniel T Gillespie. 2007. Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* 58 (2007), 35–55.

[17] Andrew G Hart and Richard L Tweedie. 2012. Convergence of invariant measures of truncation approximations to Markov processes. (2012).

[18] MHR Khouzani, Saswati Sarkar, and Eitan Altman. 2012. Maximum damage malware attack in mobile wireless networks. *IEEE/ACM Transactions on Networking* 20, 5 (2012), 1347–1360.

[19] Juan Kuntz, Philipp Thomas, Guy-Bart Stan, and Mauricio Barahona. 2019. Stationary distributions of continuous-time Markov chains: a review of theory and truncation-based approximations. *arXiv preprint arXiv:1909.05794* (2019).

[20] Thomas G Kurtz. 1978. Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications* 6, 3 (1978), 223–240.

[21] Xin Liu, Lei Ying, et al. 2020. Beyond Scaling: Calculable Error Bounds of the Power-of-Two-Choices Mean-Field Model in Heavy-Traffic. *arXiv preprint arXiv:2012.06613* (2020).

[22] Yuanyuan Liu and Wendi Li. 2018. Error bounds for augmented truncation approximations of Markov chains via the perturbation method. *Advances in Applied Probability* 50, 2 (2018), 645–669.

[23] Yuanyuan Liu, Wendi Li, and Hiroyuki Masuyama. 2018. Error bounds for augmented truncation approximations of continuous-time Markov chains. *Operations Research Letters* 46, 4 (2018), 409–413.

[24] Laurent Massoulié and Milan Vojnovic. 2005. Coupon replication systems. In *Proceedings of the 2005 ACM SIGMETRICS International Conference on Measurement and modeling of computer systems*. 2–13.

[25] Hiroyuki Masuyama. 2017. Error bounds for last-column-block-augmented truncations of block-structured Markov chains. *Journal of the Operations Research Society of Japan* 60, 3 (2017), 271–320.

[26] Wouter Minnebo and Benny Van Houdt. 2013. A fair comparison of pull and push strategies in large distributed networks. *IEEE/ACM Transactions on Networking* 22, 3 (2013), 996–1006.

[27] Michael Mitzenmacher. 2001. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems* 12, 10 (2001), 1094–1104.

[28] Brian Munsky and Mustafa Khammash. 2006. The finite state projection algorithm for the solution of the chemical master equation. *The Journal of Chemical Physics* 124, 4 (2006), 044104.

[29] Abhay K Parekh and Robert G Gallager. 1993. A generalized processor sharing approach to flow control in integrated services networks: the single-node case. *IEEE/ACM Transactions on Networking* 3 (1993), 344–357.

[30] Abhay K Parekh and Robert G Gallager. 1994. A generalized processor sharing approach to flow control in integrated services networks: The multiple node case. *IEEE/ACM Transactions on Networking* 2, 2 (1994), 137–150.

[31] E Seneta. 1967. Finite approximations to infinite non-negative matrices. In *Proc. Camb. Phil. Soc*, Vol. 63. 983.

[32] E Seneta. 1968. Finite approximations to infinite non-negative matrices, II: refinements and applications. In *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 64. Cambridge University Press, 465–470.

[33] Abhyudai Singh and Joao Pedro Hespanha. 2006. Lognormal moment closures for biochemical reactions. In *IEEE 45th Conference on Decision and Control (CDC)*. 2063–2068.
[34] John N Tsitsiklis and Kuang Xu. 2011. On the power of (even a little) centralization in distributed processing. *ACM SIGMETRICS Performance Evaluation Review* 39, 1 (2011), 121–132.
[35] Richard L Tweedie. 1971. Truncation procedures for non-negative matrices. *Journal of Applied Probability* 8, 2 (1971), 311–320.
[36] Richard L Tweedie. 1973. The calculation of limit probabilities for denumerable Markov processes from infinitesimal properties. *Journal of Applied Probability* (1973), 84–99.
[37] Benny Van Houdt. 2013. A mean field model for a class of garbage collection algorithms in flash-based solid state drives. *ACM SIGMETRICS Performance Evaluation Review* 41, 1 (2013), 191–202.
[38] N. G. Van Kampen. 2007. *Stochastic Processes in Physics and Chemistry* (3rd ed.). Elsevier.
[39] Thirupathaiah Vasantam and Ravi R Mazumdar. 2019. Fluctuations Around the Mean-Field for a Large Scale Erlang Loss System Under the SQ (d) Load Balancing. In *2019 31st International Teletraffic Congress (ITC 31)*. IEEE, 1–9.
[40] Qiaomin Xie, Xiaobo Dong, Yi Lu, and Rayadurgam Srikant. 2015. Power of d choices for large-scale bin packing: A loss model. *ACM SIGMETRICS Performance Evaluation Review* 43, 1 (2015), 321–334.
[41] Lei Ying. 2016. On the approximation error of mean-field models. *ACM SIGMETRICS Performance Evaluation Review* 44, 1 (2016), 285–297.
[42] Lei Ying. 2017. Stein's method for mean field approximations in light and heavy traffic regimes. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 1, 1 (2017), 1–27.
[43] Lulai Zhu, Giuliano Casale, and Iker Perez. 2020. Fluid approximation of closed queueing networks with discriminatory processor sharing. *Performance Evaluation* 139 (2020), 102094.

A PROOFS OF AUXILIARY RESULTS

A.1 Proof of Proposition 3.1

Proposition A.1. *Assume that $x \in \partial\mathcal{T}_y^{(l,n)}$ and for $X_y^{(n)}$ transition l is redirected to state $x^* \in \mathcal{T}_y^{(n)}$, defined by BP as in (7). Then, for any $w \in \mathbb{Z}^m$, for the process $X_{y+w}^{(n)}$, the transition from $x+w \in \partial\mathcal{T}_{y+w}^{(l,n)}$ is redirected to $x^* + w \in \mathcal{T}_{y+w}^{(n)}$.*

Proof. Suppose that for a truncation $\mathcal{T}_y^{(n)}$ and a state $x \in \partial\mathcal{T}_y^{(l,n)}$, applying Boundary Projection the transition l is redirected to x^* given by (7). Let $w \in \mathbb{Z}^m$. Then, $x+w \in \partial\mathcal{T}_{y+w}^{(l,n)}$ and the transition is redirected to $(x+w)^*$ such that:

$$\text{if } (x+w+l)_i > (x+w)_i \Rightarrow \min(y_i+w_i+n_i, x_i+w_i+l_i) = w_i + \min(y_i+n_i, x_i+l_i) = w_i + x_i^*;$$
$$\text{if } (x+w+l)_i < (x+w)_i \Rightarrow \max(y_i+w_i, x_i+w_i+l_i) = w_i + \max(y_i, x_i+l_i) = w_i + x_i^*.$$

□

A.2 Proof of Proposition 4.1

Proposition A.2. $\Pi^{(n)}(x,y) = \Pi^{(n)}(x-y,0) + y$ *for all* $x, y \in \mathcal{S}$.

Proof. Proceeding by cases:

$$\Pi_i^{(n)}(x,y) = x_i - n_i \Leftrightarrow x_i > y_i + n_i \Leftrightarrow x_i - y_i > n_i \Leftrightarrow \Pi_i^{(n)}(x-y,0) = x_i - y_i - n_i;$$
$$\Pi_i^{(n)}(x,y) = x_i \Leftrightarrow x_i < y_i \Leftrightarrow x_i - y_i < 0 \Leftrightarrow \Pi_i^{(n)}(x-y,0) = x_i - y_i;$$
$$\Pi_i^{(n)}(x,y) = y_i \Leftrightarrow y_i \le x_i \le y_i + n_i \Leftrightarrow 0 < x_i - y_i < n_i \Leftrightarrow \Pi_i^{(n)}(x-y,0) = 0.$$

□

A.3 Proof of Proposition 4.2

Proposition A.3. *The following two propositions hold:*

- *For $n = 0$, the DBP equations (15)-(16) yield the mean-field approximation of $X(t)$.*

- *For $n = \infty$, the DBP equations (15)-(16) yield the master equation of $X(t)$.*

Proof. Let us start from the case $n = 0$. For $n = 0$ we have $\mathcal{T}_0^{(n)} = \{0\}$ so every truncation has a single state. By definition the generator of a process with a single state is 0, so in (16) we have:

$$\frac{dP^0}{dt} = 0 \Rightarrow P^0(t) = P^0(0) = 1 \quad \forall t.$$

Substituting this in (15):

$$\frac{dY^0}{dt} = \sum_{l \in \mathcal{L}} \mathcal{Y}^{(l,0)}(0) f_l(Y^0(t)) = \sum_{l \in \mathcal{L}} l f_l(Y^0(t)).$$

Let us now consider $n = \infty$. For $n = \infty$ we have $\mathcal{T}_0^{\infty} = \mathcal{S}$ so the truncation covers the whole state space. Since there is no state outside the considered truncation, we do not need any augmentation and (16) for $n = \infty$ is exactly the Master Equation. Moreover $\partial \mathcal{T}_0^{(l,\infty)} = \emptyset$ for any l, since no transitions exits the observed state so we have:

$$\frac{dY^\infty}{dt} = 0 \Rightarrow Y^\infty(t) = Y^\infty(0) = 0 \quad \forall t.$$

□

B REFINED ESTIMATIONS FOR EGALITARIAN PROCESSOR SHARING

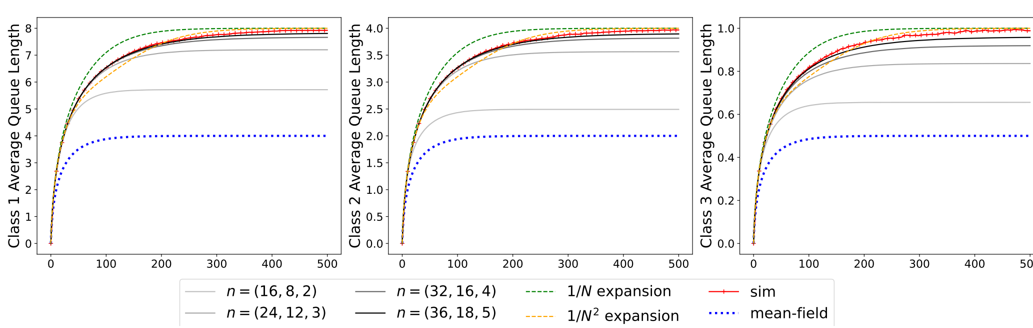

Fig. 10. Numerical results for the egalitarian processor sharing queuing system with 3 classes of customers.

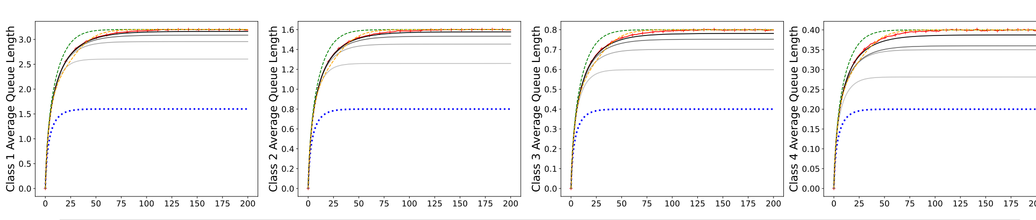

Fig. 11. Numerical results for the egalitarian processor sharing queuing system with 4 classes of customers.

C PARAMETERS FOR COXIAN DISTRIBUTION

Table 6. Parameters for Coxian service time distribution with larger number of phases.

K	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	μ_8	μ_9	μ_{10}
5	1.5136	0.5045	0.5045	0.1261	0.1261	-	-	-	-	-
10	1.8809	0.6270	0.6270	0.3135	0.3135	0.2821	0.1881	0.1567	0.0784	0.0784
K	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	
5	0.10	0.10	0.85	0.80	-	-	-	-	-	
10	0.20	0.20	0.20	0.50	0.50	0.80	0.80	0.80	0.80	

Received February 2021; revised April 2021; accepted April 2021