Batch Control of the Master Equation: A Linear Programming Approach

Jorge Gonçalves† and Nuno C Martins‡

This paper proposes a new framework for the design of finite-valued, continuous-time and homogeneous Markov processes. In particular, we propose a paradigm for selecting an optimal matrix within a pre-specified pencil of stochastic matrices. Given any stochastic matrix specifying the time-evolution of a continuous-time homogeneous Markov chain, we propose a class of figures of merit that captures the long-term evolution of any statistical moment. We show that optimization with respect to the aforementioned class of cost functions is tractable via dualization and linear programming methods. In addition, we suggest how this approach can be used as a tool for the optimal design of the master equation. Our results are applied to illustrative examples.

1 Introduction

A master equation is a probabilistic differential equation that describes a system defined by discrete configurations, i.e., it represents the continuous-time evolution of a probability distribution, and is characterized by a linear structure. A master equation, associated with a discrete-state continuous-time Markov Process [5], can represent a large number of dynamical stochastic models from queueing theory to biology. It is useful in describing systems at small length scales when continuum
assumptions break down, particularly when fluctuations are important. A master equation can be equivalently represented as a linear system, and can be finite or infinite-dimensional.

In most cases a master equation cannot be solved directly and simulations are often conducted via Monte Carlo methods using Gillespie’s stochastic algorithm [3, 4]. Although Gillespie’s stochastic stimulation can produce detailed realizations of stochastically evolving systems, the method easily becomes computationally expensive. In these cases it is often necessary to sacrifice some of the precision of the simulation for faster, yet approximate methods such as time-leaping and system-partitioning methods. Several methods have also been proposed to directly solve or approximate a master equation (see for example [2, 6, 8, 9]).

In contrast with analysis and simulation, this paper considers the problem of designing optimal parameters to minimize some cost function of the distribution (e.g. variance). This is a relevant problem when designing parameters for technological systems. In biology it can help understand some of the values of parameters that have evolved over billions of years, or help choose parameters when designing a new system in synthetic biology. We show that for finite-dimensional master equations, finding design variables appropriately so that the master equation conforms with pre-defined performance objectives can be written as a linear program which can be solved efficiently using available software.

This paper is organized as follows. The next section defines some useful concepts and introduces master equations. The problem formulation and the main technical result can be found in section 3. Then section 4 gives some illustrative examples.

We adopt the following notation:

- Given a positive integer $q$, a real vector in $\mathbb{R}^q$ is represented using bold small caps letters, such as $z \in \mathbb{R}^q$. In order to simplify notation, the components of a vector $z \in \mathbb{R}^q$ are numbered from 0 to $q-1$ using the following column vector convention: $z = (z_0, \ldots, z_{q-1})'$, where $z'$ denotes transpose of $z$.

- Real matrices are represented using large caps bold face letters, such as $Z \in \mathbb{R}^{q \times q}$.

2 Mathematical preliminaries and definitions

Consider $l$ random variables $\{X_i\}_{i=1}^l$, where each $X_i$ takes values on the set $\{0, \ldots, m_i\}$, where $\{m_i\}_{i=1}^l$ are non-negative integers. In addition, we define another random variable $X$ as follows:

$$X \overset{\text{def}}{=} (X_1, \ldots, X_l)$$

Here we use $\mathcal{M}_l$ to denote the set of all possible values that $X$ can take, or more explicitly:

$$\mathcal{M}_l \overset{\text{def}}{=} \{0, \ldots, m_1\} \times \ldots \times \{0, \ldots, m_l\}$$
The following represents the set of possible probability mass vectors of $X$:

$$\mathbb{P}^m \overset{\text{def}}{=} \{ \mathbf{v} \in \mathbb{R}^{m+1}_{\geq 0} : \sum_{i=0}^{m} v_i = 1 \}$$

where $m \overset{\text{def}}{=} \prod_{i=1}^{l} (m_i + 1) - 1$ is the cardinality of $M_l$.

Given a bijective map $S : M_l \rightarrow \{0, \ldots, m\}$, the elements of the set $\mathbb{P}^m$ can be viewed as probability mass functions of $X$. In particular, for any vector $\mathbf{v}$ in the set $\mathbb{P}^m$, we assume that $\text{Prob}(X = x) = \mathbf{v}_{S(x)}$ with $x \in M_l$. As such, in our formulation, the probability mass function of $X$ is represented in a columnar way, as elements of $\mathbb{P}^m$.

Given a non-negative integer $m$, a probability mass vector $\mathbf{v} \in \mathbb{P}^m$ and a function $F : M_l \rightarrow \mathbb{R}$, the expected value of $F(X)$ is given by:

$$E[F(X)] = \mathbf{f}' \mathbf{v}$$

where $\mathbf{f}_i \overset{\text{def}}{=} F(S^{-1}(i))$, with $i \in \{0, \ldots, m\}$.

**Example 2.1.** We can use the above representation to express the moments of a random variable. For instance, for $l = 1$ and given any non-negative integer $m$, if the vector $\mathbf{f} \in \mathbb{R}^{m+1}$ is given by $f_x = x^2$ for $x \in \{0, \ldots, m\}$ then, for any probability mass vector $\mathbf{v} \in \mathbb{P}^m$, the following holds:

$$\mathbf{f}' \mathbf{v} = E[X^2]$$

Similarly, if $f_x = |x - 3|$ then $\mathbf{f}' \mathbf{v} = E[|X - 3|]$, the expected distance between $X$ and 3.

A matrix $\mathbf{H}$ is stochastic if the sum of the elements in each of their columns is zero.

### 2.1 Matrix form of the master Equation

For simplicity, we write the time evolution of the probability mass vector of Markovian continuous-time and finite-space processes in the following compact form.

**Definition 1.** Given non-negative integers $m_1, \ldots, m_l$, $n$, a parameter vector $\mathbf{c} \in \mathbb{R}^n$ and matrices $\mathbf{H}_0, \ldots, \mathbf{H}_n \in \mathbb{R}^{(m+1) \times (m+1)}$, with $m \overset{\text{def}}{=} \prod_{i=1}^{l} (m_i + 1) - 1$,
the following is a state-space parameterized form of the master Equation:

$$\frac{d}{dt} \mathbf{p}(t) = \mathbf{H}(\mathbf{c}) \mathbf{p}(t), \quad t \geq 0, \quad \mathbf{p}(0) \in \mathbb{P}^m$$

where

$$\mathbf{H}(\mathbf{c}) = \mathbf{H}_0 + \sum_{k=1}^{n} c_k \mathbf{H}_k$$
and \( \mathbf{p}(t) \) is the probability mass vector of a random variable taking values in the set \( \mathcal{M}_1 \) at time \( t \).

Markov processes can also be defined in a countable infinite space but these will not be considered here.

**Example 2.2 (Birth and death processes).** Let \( l = 1 \). Given positive real constants \( c_1 \) and \( c_2 \), and a non-negative integer \( m \) representing the maximum number of individuals, consider the following birth and death process dynamics:

\[
\emptyset \xrightarrow{c_1}{c_2} X
\]

Hence, a “birth” occurs according to a Poisson process with a probability per unit time \( c_1 \) and a “death” with probability per unit time proportional to \( X \), i.e., \( c_2 X \). Alternatively, the process can be visualized as

\[
0 \xrightarrow{c_1}{c_2} 1 \xrightarrow{c_1}{2c_2} 2 \cdots X \xrightarrow{c_1}{c_2X} X \xrightarrow{c_1}{c_2(X+1)} X + 1 \cdots m - 1 \xrightarrow{c_1}{c_2m} m
\]

Here we assume the maximum value for \( X \) is \( m \).

The master equation can be found directly from the Markov process:

\[
\frac{d}{dt} p_x(t) = \begin{cases} -c_1 p_0(t) + c_2 p_1(t) & \text{if } x = 0 \\ c_1 p_{x-1}(t) - (c_1 + c_2 x) p_x(t) + c_2 (x+1) p_{x+1}(t) & \text{if } x \in \{1, \ldots, m-1\} \\ c_1 p_{m-1}(t) - c_2 m p_m(t) & \text{if } x = m \end{cases}
\]

for all \( t \in \mathbb{R}_{\geq 0} \), and where \( p_x(t) \) represents the probability of having \( x \) individuals at time \( t \). Hence, in this case \( H_0 = 0 \).

\[
H_1 = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 1 & -1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \end{bmatrix}, \quad H_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ 0 & -2 & 3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & -(m-1) & m \\ 0 & \cdots & 0 & 0 & -m \end{bmatrix}
\]

**Example 2.3 (Schlogl model).** Let \( l = 1 \). More complex interactions can be considered. For example, the Schlogl model is described by the following chemical reaction network \([5]\)

\[
A + 2X \xrightarrow{c_1AX}{c_2X^2} 3X \quad B \xrightarrow{c_2B}{c_3B} X
\]

where the concentrations \( A \) and \( B \) are kept constant (buffered). For simplicity of
notation, let $c_1 = A\bar{c}_1$ and $c_3 = B\bar{c}_3$. In this case, the master equation is given by
\[
\frac{d}{dt}p_x(t) = \left\{
\begin{array}{ll}
-c_0 p_0(t) + c_1 p_1(t) & \text{if } x = 0 \\
-c_0 p_0(t) - (c_4 + c_3)p_2(t) + 2c_0 p_2(t) & \text{if } x = 1 \\
(c_1^{(x-1)(x-2)} + c_3) p_{x-1}(t) \\
- \left( c_1^{(x-1)} + c_2^{(x-1)(x-2)} + c_3 + c_4 x \right) p_x(t) \\
+ \left( c_2^{(x+1)(x-1)} + c_4 (x+1) \right) p_{x+1}(t) & \text{if } x \in \{2, \ldots, m-1\} \\
(c_1^{(m-1)(m-2)} + c_3) p_{m-1}(t) \\
- \left( c_2^{m(m-1)(m-2)} + c_4 m \right) p_m(t) & \text{if } i = m
\end{array}
\right.
\]

(3)

Just as above, from these matrices $H_1$ to $H_4$ can be easily constructed (with $H_0 = 0$).

Other more complex master equations can be considered, including multiple species.

**Example 2.4 (Two species).** Let $l = 2$. Consider two species $X_1$ and $X_2$, where $X_1$ activates $X_2$, $X_2$ inhibits $X_1$, and each is degraded over time according to the chemical reaction network

\[
A \xrightarrow{c_1 h(X_2)} X_1 \quad X_1 \xrightarrow{c_2} \emptyset \\
X_1 \xrightarrow{c_2 X_2} X_2 \quad X_2 \xrightarrow{c_2} \emptyset
\]

or, as an alternative representation

\[
\begin{array}{c}
A \\
| \quad | \quad | \\
X_1 \xrightarrow{c_2} X_2
\end{array}
\]

where the concentration $A$ is kept constant (buffered) and $h(\cdot)$ is a Hill-equation given by $h(X_2) = A/(1 + 0.04 X_2^2)$. In this case the master equation (not shown) is a linear differential equation in the variables $p_{x_1 x_2}(t)$. The matrices $H_1$ to $H_2$ can again be constructed (with $H_0 = 0$).

### 3 Problem Formulation and Main Technical Result

In our formulation, the dynamic matrix $H(c)$ of the master equation (1) is an affine function of the entries of $c$. We are interested in cases where the entries of $c$ are design variables, such as the constants $c_1$ and $c_2$ in (2), and $c_1$ to $c_4$ in (3), which must be selected appropriately so that the master equation conforms with pre-defined performance objectives (see Section 4). The main focus of this paper is to present a method for computing the optimal $c$ with respect to a family of performance measures.
3.1 Definition of a new performance metric, and discussion of its suitability

In this section, we start by defining a family of performance measures for the master equation, followed by a brief description of its main properties. Subsequently, we formulate an optimization paradigm, with respect to the vector $c$, which can be solved using linear programming techniques [1].

**Definition 2.** Given a non-negative integer $m$, a positive real number $\epsilon$, a real vector $f \in \mathbb{R}_{\geq 0}^{m+1}$, and a stochastic matrix $H \in \mathbb{R}^{(m+1) \times (m+1)}$, we define the following map:

$$E_f(H, \epsilon) \overset{\text{def}}{=} \max_{v \in \mathbb{P}^m_H, \epsilon} f'v$$

where the set $\mathbb{P}^m_{H, \epsilon} \subset \mathbb{P}^m$ is given by:

$$\mathbb{P}^m_{H, \epsilon} \overset{\text{def}}{=} \{ v \in \mathbb{P}^m | f'Hv \geq -\epsilon f'v \}$$

We proceed with the following Proposition that constitutes one of the main motivations for adopting $E_f(H, \epsilon)$ as a performance metric.

**Proposition 3.** Given a non-negative integer $m$, a real vector $f \in \mathbb{R}_{\geq 0}^m$, and a stochastic matrix $H \in \mathbb{R}^{(m+1) \times (m+1)}$, the following holds:

$$\limsup_{t \to \infty} \sigma(t) \leq E_f(H, \epsilon), \quad \epsilon > 0, \quad p(0) \in \mathbb{P}^m$$

where $\sigma(t) \overset{\text{def}}{=} f'p(t)$ and $p(t)$ obeys the following master equation:

$$\frac{d}{dt}p(t) = Hp(t), \quad t \geq 0, \quad p(0) \in \mathbb{P}^m$$

As we illustrated in Example 2.1 and preceding discussion, any metric based on the expected value of a function of $X(t)$ can be represented via the inner product between $p(t)$ and an appropriately chosen vector $f$. This fact and Proposition 3 unveil the significance of $E_f(H, \epsilon)$ as an upper-bound to the time limit of performance metrics, provided that they are based on the statistical expectation of functions of $X(t)$. Further properties of $E_f(H, \epsilon)$ are discussed in Section 3.3.

**Proof.** (of Proposition 3) We start by defining the following set:

$$\mathbb{A}^m_{H, \epsilon} \overset{\text{def}}{=} \{ v \in \mathbb{P}^m | f'v \leq E_f(H, \epsilon) \}$$

The proof follows from the following properties of the set $\mathbb{A}^m_{H, \epsilon}$:

$$p(t^*) \notin \mathbb{A}^m_{H, \epsilon} \implies \left. \frac{d}{dt}f'p(t) \right|_{t=t^*} < -\epsilon f'p(t^*), \quad t^* \geq 0$$
\[ p(t^*) \in A^m_{H, \epsilon} \implies p(t^* + \delta) \in A^m_{H, \epsilon}, \quad \delta > 0, t^* \geq 0 \]

where we used the fact that, from (6), \( \frac{d}{dt} f'(t^*) = \Gamma H p(t^*) \).

It also follows from (7) that all trajectories of \( p(t) \) will converge exponentially to the set \( A^m_{H, \epsilon} \).

### 3.2 Optimization paradigm and its solution using linear programming

Motivated by the discussion in sections 3.1 and 3.3, we investigate the following optimization paradigm:

**Optimization paradigm:** Given non-negative integers \( m \) and \( n \), a positive real \( \epsilon \), stochastic matrices \( H_0, \ldots, H_n \in \mathbb{R}^{(m+1)\times(m+1)} \), a vector \( f \in \mathbb{R}^{m+1} \), a lower bound vector \( c \in \mathbb{R}^n \) and an upper bound vector \( \bar{c} \in \mathbb{R}^n \), satisfying \( c_k \leq c_k \leq \bar{c}_k \), we wish to solve the following optimization problem:

\[
J^* = \min_{c \in \mathbb{R}^n} \mathcal{E}_f(H(c), \epsilon) \tag{8}
\]

subject to:

\[
\begin{align*}
H(c) &= H_0 + \sum_{k=1}^{n} c_k H_k \\
c_k &\leq c_k \leq \bar{c}_k, \quad k \in \{1, \ldots, n\}
\end{align*} \tag{9}
\]

**Remark 3.1** (Further remarks on the comparability between \( \mathcal{E}_f(H, \epsilon) \) and \( \lim_{t \to \infty} \sup \Gamma p(t) \)). It follows from (5) that our optimization paradigm is guaranteed to minimize an upper-bound to \( \limsup_{t \to \infty} \Gamma p(t) \). However, in general, equality in (5) does not hold because the definition of \( \mathcal{E}_f(H, \epsilon) \) imposes exponential rates of convergence of \( -\epsilon \) (see (7)). The fact that the gap in (5) is non-decreasing with \( \epsilon \) indicates that the conservativeness, of such an upper-bound minimization, might be reduced by selecting smaller \( \epsilon \), at the expense of less guarantees in terms of the convergence rate. Further properties of \( \mathcal{E}_f(H, \epsilon) \) are discussed in Section 3.3.

The following Theorem shows that our optimization paradigm can be efficiently solved by means of a linear program.

**Theorem 4.** Consider that the parameters defining the optimization paradigm (8)-(9) are given. A solution to the optimization paradigm described by (8)-(9) is given by:

\[
c_k^* = \begin{cases} 
\frac{d_k^*}{\lambda^*} & \text{if } \lambda^* \neq 0 \\
\text{any } c_k^* \text{ satisfying } c_k \leq c_k^* \leq \bar{c}_k & \text{otherwise}
\end{cases}, \quad k \in \{1, \ldots, n\} \tag{10}
\]

where \( d_1^*, \ldots, d_n^* \) and \( \lambda^* \) are an optimal solution of the following linear program:

\[
J^* = \min_{\lambda \geq 0, d \in \mathbb{R}^n, \gamma \geq 0} \gamma \tag{11}
\]
subject to
\[ f' \left[ I + \lambda (\epsilon I + H_0) + \sum_{k=1}^{n} d_k H_k \right] E_{i,i} \leq \gamma, \quad i \in \{0, \ldots, m\} \tag{12} \]
\[ \lambda \lambda_k \leq d_k \leq \lambda \bar{c}_k, \quad k \in \{1, \ldots, n\} \]

where, for each \( i \) in the set \( \{1, \ldots, m + 1\} \), \( E_{i,i-1} \) is a column vector in \( \mathbb{R}^{m+1} \) representing the \( i \)th canonical basis of \( \mathbb{R}^m \), i.e., \([E_{0,0} \cdots E_{m,m}]\) is the identity matrix in \( \mathbb{R}^{(m+1) \times (m+1)} \).

**Proof. (Case I):** We start by proving the result under the assumption that \( J^* < \bar{f} \), where \( \bar{f} \equiv \max_{i \in \{0, \ldots, m\}} f_i \). From Definition 2, we can re-cast the optimization paradigm described by (8)-(9) as follows:
\[ \min \max c_{i \in \{0, \ldots, m\}} f' v \tag{13} \]
subject to (9) and to the following inequality obtained from (4):
\[ f'H(c)v + \epsilon f'v \geq 0 \tag{14} \]

We start by noticing that, by introducing a Lagrange multiplier \( \lambda \), the optimization paradigm described by (13), (9) and (14) can be equivalently written as:
\[ \min c \max v \in \mathbb{P}^m \min f'v + \lambda \left( f' \left( H_0 + \sum_{k=1}^{n} c_k H_k + \epsilon I \right) v \right) \tag{15} \]
subject to: \( c_k \leq c_k \leq \bar{c}_k \), \( k \in \{1, \ldots, n\} \)
\[ \tag{16} \]
Since the Lagrangian in (15) is convex in \( \lambda \) and concave in \( v \), the duality gap is zero [7] and we can exchange the order between \( \max_{v \in \mathbb{P}^m} \) and \( \min_{\lambda \geq 0} \). After the aforementioned order exchange and by using the substitution \( d_k = \lambda \bar{c}_k \), we get that (15)-(16) is equivalent to:
\[ \min d \in \mathbb{R}^n \min_{\lambda \geq 0} \max_{v \in \mathbb{P}^m} \left( f' \left( \lambda H_0 + \sum_{k=1}^{n} d_k H_k + (\lambda + 1)I \right) v \right) \tag{17} \]
subject to: \( \lambda \lambda_k \leq d_k \leq \lambda \bar{c}_k \), \( k \in \{1, \ldots, n\} \)
\[ \tag{18} \]
Notice that the maximizing \( v \) in (17) belongs to the set \( \{E_{1,1}, \ldots, E_{m,m}\} \), where, for each \( i \) in the set \( \{1, \ldots, m + 1\} \), \( E_{i,i-1} \) is a column vector in \( \mathbb{R}^{m+1} \) representing the \( i \)th canonical basis of \( \mathbb{R}^m \), i.e., \([E_{0,0} \cdots E_{m,m}]\) is the identity matrix in \( \mathbb{R}^{(m+1) \times (m+1)} \). Hence, we can re-write (17) as follows:
\[ \min d \in \mathbb{R}^n \min_{\lambda \geq 0} \max_{i \in \{0, \ldots, m\}} f' \left[ I + \lambda (\epsilon I + H_0) + \sum_{k=1}^{n} d_k H_k \right] E_{i,i} \tag{19} \]
subject to: \( \lambda \lambda_k \leq d_k \leq \lambda \bar{c}_k \), \( k \in \{1, \ldots, n\} \)
\[ \tag{20} \]
Notice that since \( J^* < \bar{f} \) then the constraint (14) will be active at any optimum, and that implies that \( \lambda^* \) must be nonzero. The proof for the case \( J^* < \bar{f} \) is concluded by noticing that (19)-(20) is equivalent to (11)-(12).
(Case II): Now, assume that \( J^* = \bar{f} \). (Note: From the analysis above, we conclude that \( \lambda^* = 0 \) implies that \( J^* = \bar{f} \) holds. On the other hand, \( \lambda^* \neq 0 \) does not imply that \( J^* = \bar{f} \).) Notice, from Definition 2, that \( \mathcal{E}_F(H(c), \epsilon) \) in (8) satisfies \( \mathcal{E}_F(H(c), \epsilon) \leq \bar{f} \). Hence, if \( J^* = \bar{f} \) holds then the optimization in (8)-(9) is ineffectual, and any choice of \( c \) within the specified bounds is optimal.

(proof of (10)): In order to finalize the proof of the Theorem, we proceed to deriving the form of the optimal \( c^* \) as described in (10). If \( \lambda^* \neq 0 \) then the solution can be obtained by reversing the substitution \( d_k^* = c_k^* \lambda^* \), i.e., \( c_k^* = \frac{d_k^*}{\lambda^*} \). On the other hand, if \( \lambda^* = 0 \) then, from (Case II), we conclude that \( J^* = \bar{f} \) holds and any selection of \( c^* \), within the required bounds, is optimal. \( \square \)

3.3 Additional properties of \( \mathcal{E}_F(H, \epsilon) \).

In this sub-Section, we explain why \( \mathcal{E}_F(H, \epsilon) \) is the smallest real constant that can be used to characterize the exponentially attractive set \( A_{H, \epsilon} \), via a direct comparison with the performance metric \( f'(p(t)) \). In the following Remark, we re-write \( \mathcal{E}_F(H, \epsilon) \) in a form that will facilitate the presentation of Remark 3.3, where we present the main conclusion of this sub-Section.

Remark 3.2. Given an integer \( m \) and a vector \( f \) in the set \( \mathbb{R}^{m+1}_0 \), let \( \bar{f} \) denote the maximum among the entries of \( f \), i.e., \( \bar{f} \overset{\text{def}}{=} \max_{k \in \{0, \ldots, m\}} f_k \). Given \( \epsilon > 0 \) and a matrix \( H \) in \( \mathbb{R}^{(m+1)\times(m+1)} \), if \( \mathcal{E}_F(H, \epsilon) < \bar{f} \) then the following is a consequence of Definition 2:

\[
\mathcal{E}_F(H, \epsilon) = \inf \left\{ \gamma \in \mathbb{R} \mid (f'v > \gamma) \implies (f'Hv < -\epsilon f'v), \; v \in \mathbb{P}^m \right\}
\]

Remark 3.3. Consider a continuous-time Markovian process \( X(t) \) taking values on the set \( \{0, \ldots, m\} \) and whose probability mass function obeys the following (master equation) dynamics:

\[
\frac{d}{dt} p(t) = H p(t), \; t \geq 0, \; p(t) \in \mathbb{P}^m
\]

where \( p_i(t) \overset{\text{def}}{=} \text{Prob}(X(t) = i) \). From Remark 3.2, the following holds:

\[
\mathcal{E}_F(H, \epsilon) = \inf \left\{ \gamma \in \mathbb{R} \mid (\sigma(t^*) > \gamma) \implies \left( \frac{d}{dt} \sigma(t) \right)_{t = t^*} < -\epsilon \sigma(t^*), \; p(t^*) \in \mathbb{P}^m, t^* \geq 0 \right\}
\]

(21)

where we define \( \sigma(t) \overset{\text{def}}{=} f'p(t) \). As such, if \( p(t^*) \) is such that \( \sigma(t^*) > \mathcal{E}_F(H, \epsilon) \) holds then, at time \( t^* \), \( \sigma(t^*) \) will decrease with an exponentially rate of \( -\epsilon \). Hence, we conclude that \( \mathcal{E}_F(H, \epsilon) \) is the smallest number \( \gamma \) for which instantaneous exponential decrease can be verified based on the inequality \( \sigma(t^*) > \gamma \), with respect to all possible \( p(t^*) \).

\(^1\)Recall that, as illustrated in Example 2.1, the expectation of any function of \( X(t) \) is a linear map of \( p(t) \).
4 Examples

Next we consider again the three examples introduced in section 2.1.

4.1 Birth and death processes

Let \( m = 100 \) and define \( f_x = (x - 35)^2 \). This will minimize the variance of the Markov process with an average at 35. The results can be seen in Fig 1.

![Figure 1](image)

*Figure 1. Left: log-log plot for different values of \( \epsilon \) versus the error = \( \gamma^* - f v_\infty \), where \( \gamma^* \) is the optimal solution in (11) and \( v_\infty \) is the stationary solution of (1); the trade-offs are clear: the faster the convergence rate, the more conservative is the problem. Right: stationary distribution (solid) and \( f_x \) (dotted), scaled for comparison.*

4.2 Schlogl model

The Schlogl model has four degrees of freedom \( c_1 \) to \( c_4 \). The left of Fig 2 shows the stationary distribution of the Schlogl model with \( c_1 = 0.01, c_2 = 0.000085, c_3 = 500 \) and \( c_4 = 0.5 \), obtained from the optimization with \( f_x = (x - 450)^2 \) and \( \epsilon = 0.1 \).

With a different optimization function, a bimodal distribution can be obtained with \( c_1 = 0.0120217, c_2 = 0.00005, c_3 = 50, \) and \( c_4 = 1.1181 \), and with \( \epsilon = 1 \).

4.3 Two species

Consider two degrees of freedom \( c_1 \) and \( c_2 \). The first Fig 3 shows the desired optimization function \( f_x = (x_1 - 10)^2 + (x_2 - 10)^2 \). The other two figures show the resulting stationary distribution with \( c_1 = 20 \) and \( c_2 = 1.02 \).
Figure 2. Left: stationary distribution (solid) and \( f_x \) (dotted), scaled for direct comparison. Right: bimodal stationary distribution (solid) and \( f_x \) (dotted).

Figure 3. Top left: optimization function \( f_x \). Top right: stationary distribution. Bottom: stationary distribution on the plane \( x_1 = x_2 \).
Bibliography


