Estimation of state-transition probability matrices in asynchronous population Markov processes

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5531431">http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5531431</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Institute of Electrical and Electronics Engineers (IEEE)</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/76753">http://hdl.handle.net/1721.1/76753</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>

Waleed A. Farahat and H. Harry Asada

Abstract—We address the problem of estimating the probability transition matrix of an asynchronous vector Markov process from aggregate (longitudinal) population observations. This problem is motivated by estimating phenotypic state transitions probabilities in populations of biological cells, but can be extended to multiple contexts of populations of Markovian agents. We adopt a Bayesian estimation approach, which can be computationally expensive if exact marginalization is employed. To compute the posterior estimates efficiently, we use Monte Carlo simulations coupled with Gibb’s sampling techniques that explicitly incorporate sampling constraints from the desired distributions. Such sampling techniques can attain significant computational advantages. Illustration of the algorithm is provided via simulation examples.

I. INTRODUCTION

Many dynamic systems consisting of a population of independent agents can be described by population Markov processes (also known as vector Markov processes [1]). In these systems (illustrated in Figure 1), the discrete state of each agent evolves according to a Markov process, that is characterized by a probability transition matrix $P$. The objective of this work is to develop algorithms for the estimation of $P$ from observed data. In particular, we focus our attention to the problem when such transitions are asynchronous among the population members. That is, the individual agents do not share a common clock that dictates when the transitions occur. Furthermore, the observations are not necessarily in sync with any particular observation schedule. The distinction between synchronous and asynchronous transitions is illustrated in Figure 2.

This work is directly motivated by the problem of estimating phenotypic state transition in living cells. Consider a colony of cells or bacteria residing in a petri dish. This system can be described as population Markov process, where each cell in the population undergoes a sequence of phenotypic transitions over time. These transitions are typically not deterministic, and are best described in stochastic, particularly Markovian, terms. There are numerous examples in biology for such stochastic transitions. *Escherichia coli* stochastically switch between highly proliferating and quiescent phenotypes to maintain increased resistance to stress conditions induced by antibiotics [2]. *Candida albicans*, *Salmonella typhimurium*, and other bacterial and fungal organisms exhibit high frequency, reversible phenotypic switching to enhance anti-fungal resistance [3]. Slime mold (*Dictyostelium discoideum*) exhibits migration and formation dynamics in the presence of stress signals that are released at the onset of starvation (cAMP and follic acid), causing the phenotype of the organisms to change from independent unicellular organisms, to a formation of a multicellular organism capable of survival and reproduction [4]. Human endothelial cells, stochastically switch among several states (quiescence, proliferation, migration and apoptosis – programmed cell death) in response to external and internal signaling molecules [5], [6]. While such transitions can be synchronized initially in *vitro* (e.g. via the use of cell starvation techniques to reset cell cycle clocks), the synchronicity typically decays within several transitions. Irrespective of the particular context, it is of interest to estimate the probabilities of phenotypic state-transitions from observations collected over cell population. The transition probability matrix characterizes the dynamics of cell populations succinctly, and enables answering questions pertaining to transition dynamics, steady-state distributions and other outcomes of interest.

The estimation algorithms employed depend critically on the form of observations available. At one extreme, an abundance of data can simplify the estimation problem significantly. For example, if panel data is available (i.e. observations of the state of each agent at each observation point), the estimation problem becomes straightforward. However, if only longitudinal data is available (i.e. aggregate data summarizing the total number of agents in each state), the state-transition probability matrix estimation problem becomes more involved.

In [7], we developed Bayesian estimation schemes for addressing both estimation problems (using panel and longitudinal data). A tacit assumption was that, in both formulations, the state transitions are synchronous among the population members, and are also synchronous with the observations. In this work, we relax these assumptions, and allow for asynchronicity of state transitions among population members, and with respect to the observations. Thus, here we generalized the algorithms to a wider class of problems.

We adopt a Bayesian estimation approach to this problem for three reasons. First, Bayesian estimates can give more representative results when computing expectations of nonlinear functions over random variables. This is desirable when using the model to predict future output of the system.
II. PROBLEM FORMULATION

A. Preliminaries

Consider a population of $N$ agents. Each agent $i$ assumes a state $s_i(t) \in S$ at time $t$, where $S = \{1, 2, \cdots, m\}$ is the set of labels of states that an agent can take. The evolution $i$-th agent’s states can be described by the random process $s_i = \{s_i(1), s_i(2), \cdots, s_i(T)\}$, where $T$ is the time horizon. The transitions of the entire population of cells can be described by the $N \times T$ matrix $S$

$$
S = \begin{bmatrix}
  s_1(1) & s_1(2) & \cdots & s_1(T) \\
  s_2(1) & s_2(2) & \cdots & s_2(T) \\
  \vdots & \vdots & \ddots & \vdots \\
  s_N(1) & s_N(2) & \cdots & s_N(T)
\end{bmatrix} = \begin{bmatrix}
  s_1 \\
  s_2 \\
  \vdots \\
  s_N
\end{bmatrix} \quad (1)
$$

The matrix $S$ is the most complete description of the cell transitions since it allows for tracking the time trajectory of each individual cell. In the population statistics literature, this is known as the panel data matrix [8].

A more compact representation of the transitions is to define the $m \times 1$ aggregate count vector $\theta$ that denotes that aggregate number of cells occupying each of the states.

$$
\theta(t) = \begin{bmatrix}
  \text{# of cells in state 1 at time } t \\
  \text{# of cells in state 2 at time } t \\
  \vdots \\
  \text{# of cells in state m at time } t
\end{bmatrix} \quad (2)
$$

Thus the $k$th element of $\theta(t)$, $\theta_k(t)$, is the number of cells in the population that assume state $k$ at time $t$. Note that $\sum_k \theta_k = N$. The evolution of $\theta(t)$ over time is given by the $m \times T$ matrix $\Theta$

$$
\Theta = [\theta(0), \theta(1), \cdots, \theta(T)] \quad (3)
$$

This matrix $\Theta$ is known as aggregate or longitudinal data matrix. Note that in typical situations $m << N$, i.e. there are many more agents than states, and the matrix $\Theta$ represents highly summarized data describing the system. Figure 2 illustrates the distinction between panel data and longitudinal data for both the synchronous and asynchronous cases.

Let $p_{jk}$ denote the probability of transition of from state $j$ to state $k$. For simplicity we assume that the population is homogeneous, and thus $p_{jk}$ applies to all agents, and that the transitions of each agents is independent of each other. The $m \times m$ probability transition matrix $P = [p_{jk}]$ is unknown, and is the objective of our estimation problem. Note that for $P$ to be a valid stochastic matrix, its elements have to obey basic constraints, namely that $0 \leq p_{jk} \leq 1$ and $\sum_k p_{jk} = 1$.

B. Asynchronous Transition Model

To model the asynchronous nature of the transitions, we rely on a Poisson processes description of the temporal statistics of the transitions. This description has two consequences.
First, the time at which a particular agent visits a particular state is given by an exponential distribution, characterized by the parameter \( \lambda \). Second, for a given time interval \( \Delta t \), the probability that \( n \) transitions occur within that period is given by

\[
\Pr(n|\lambda, \Delta t) = e^{-\lambda \Delta t} \frac{(\lambda \Delta t)^n}{n!}.
\]

(4)

Such a description is common for arrival rates within given time intervals.

C. Bayesian Estimation Formulation

Our problem becomes: given longitudinal and asynchronous observations \( \Theta \), estimate the probability transition matrix \( \mathbf{P} \). Since we are adopting a Bayesian framework, we arrive at probability distributions of the parameters of interest – the elements of the probability transition matrix. Therefore we recursively apply

\[
\Pr(\mathbf{P}|\mathbf{\theta}(t), \mathbf{\theta}(t+\Delta t)) = \frac{\Pr(\mathbf{\theta}(t+\Delta t)|\mathbf{P}, \mathbf{\theta}(t)) \cdot \Pr(\mathbf{P}|\mathbf{\theta}(t))}{\Pr(\mathbf{\theta}(t+\Delta t)|\mathbf{\theta}(t))}
\]

(5)

In most cases, the computation of the denominator (a proportionality constant), is not required, and we will satisfy ourselves with relative probabilistic values. Note that the priors \( \Pr(\mathbf{P}|\mathbf{\theta}(t)) \) are either given or assumed given the particular context of the problem. The key challenge in this problem, as with most Bayesian estimation problems, is to compute the likelihood function \( \Pr(\mathbf{\theta}(t+\Delta t)|\mathbf{P}, \mathbf{\theta}(t)) \).

III. ESTIMATION FORMULATION

A. State Transition Flows

The state transition flow matrix \( \mathbf{F}(t) \) is defined by \( \mathbf{F}(t) = \{ f_{jk}(t) \} \), where \( f_{jk}(t) \) is the number of cells that have transitioned from state \( j \) to state \( k \) during the interval \( (t, t+\Delta t) \). We introduced \( \mathbf{F}(t) \) as an auxiliary variable since it simplifies the problem formulation considerably.

Given a candidate state transition matrix \( \mathbf{P} \), the probability of a flow vector out of a given node is captured by the multinomial distribution:

\[
\Pr(f_{i1}(t), f_{i2}(t), \cdots f_{im}(t)|\mathbf{P}, \mathbf{\theta}(t)) = \frac{\theta_i(t) \cdot (f_{i1}(t)f_{i2}(t) \cdots f_{im}(t)) \cdot p_{i1}^{f_{i1}(t)} \cdots p_{im}^{f_{im}(t)}}{\theta_i(t)! \cdot \prod_{k=1}^{m} \frac{p_{ik}(t)}{f_{ik}(t)!}}.
\]

(6)

The probability of the entire flow matrix \( \mathbf{F}(t) \) is the joint probability of the individual flow vectors given by equation (6). Assuming independence of the flows originating out of all the nodes, the joint probability is computed by:

\[
\Pr(\mathbf{F}(t)|\mathbf{P}, \mathbf{\theta}(t)) = \prod_{i=1}^{m} \theta_i(t)! \prod_{k=1}^{m} \frac{p_{ik}^{f_{ik}(t)}}{f_{ik}(t)!}.
\]

(7)

These flow matrices have to satisfy the consistency constraints that i) the sum of flows out of a particular node is equal to the total agents in that node, and ii) the sum of flows into a node equals the number of agents in that node at the next instance. These constraints are written as

\[
\sum_{k} f_{ik}(t) = \theta_i(t)
\]

(8)

\[
\sum_{i} f_{ik}(t) = \theta_k(t + \Delta t)
\]

(9)

Furthermore, a set of inequality constraints must be satisfied, specifically that the flows out of a node are non-negative, and are bounded by the total number of agents in a given state:

\[
f_{ik}(t) \geq 0
\]

(10)

\[
f_{ik}(t) \leq \theta_i
\]

(11)

Without loss of generality, we can compactly write equality constraints (8), (9) and inequality constraints (10), (11) as

\[
\mathbf{A}(t)\mathbf{f}(t) + \mathbf{b}(t) = \mathbf{0}
\]

(12)

\[
\mathbf{C}(t)\mathbf{f}(t) + \mathbf{d}(t) \leq \mathbf{0}
\]

(13)

We denote the set of all possible flow matrices satisfying constraints (12) and (13) by \( \mathcal{F}(t) \).

B. Estimation of State Transition Probabilities \( \mathbf{P} \)

To compute the likelihood of aggregate longitudinal observations, we marginalize over all possible flow matrices satisfying the consistency constraints. Thus,

\[
\Pr(\mathbf{\theta}(t+\Delta t)|\mathbf{P}, \mathbf{\theta}(t)) = \sum_{\mathbf{F}(t) \in \mathcal{F}(t)} \Pr(\mathbf{F}(t)|\mathbf{P}, \mathbf{\theta}(t))
\]

\[
= \sum_{\mathbf{F}(t) \in \mathcal{F}(t)} \left( \prod_{i=1}^{m} \theta_i(t)! \prod_{k=1}^{m} \frac{p_{ik}^{f_{ik}(t)}}{f_{ik}(t)!} \right)
\]

(14)

Equation (14), when substituted in equation (5) becomes the basis for our estimation scheme.

IV. COMPUTATIONAL ISSUES

The computation of the posterior distribution in eq. (14) by exhaustive inference is computationally expensive. This is primarily because the set \( \mathcal{F}(t) \) is typically of a large dimensionality \( (m^2) \), and consequently exact summation occurs over a large number of elements. To attain computational tractability, we apply the Monte Carlo principle to compute this summation approximately efficiently. We rely on two key approximations:

1) Approximate the multinomial distribution (eq. (6)) describing the likelihood functions by an equivalent multi-dimensional Gaussian distribution, implying that

\[
\Pr(\mathbf{F}(t)|\mathbf{P}, \mathbf{\theta}(t)) \approx \mathcal{N}(\mu_F, \Sigma_F)
\]

2) Apply Monte Carlo methods to evaluate the summation over the set \( \mathcal{F} \) efficiently by relying on appropriate sampling techniques. We implement this via Gibb’s sampling from the approximated Gaussian probability distributions while satisfying equality constraints.
(12) and inequality constraints (13). This implies that
\[ \sum_{F \in \mathcal{F}} \mathcal{N}(\mu_F, \Sigma_F) \approx \sum_{F_{\text{samples}} \in \mathcal{F}} \mathcal{N}(\mu_F, \Sigma_F). \]
Both of these approximations are described next in detail.

A. Gaussian Approximation to the Multinomial Likelihood Function

The likelihood functions for the flow transition matrices in eq. (7) are multinomial distributions. These functions can be readily approximated by multivariate Gaussian distributions by matching the means and the covariances appropriately [9]. To simplify the ensuing expressions, it is useful to rewrite the \( m \times m \) matrix quantities \( F \) and \( P \) as \( m^2 \times 1 \) vectors \( f \) and \( p \) by stacking the matrix elements column wise.

The expectation of a particular flow path \( f_{ik} \) is given by
\[ \mu_{f_{ik}} = E[f_{ik}] = p_{ik}\theta_i. \]
Therefore the mean of Gaussian approximation is given by
\[ \mu_f = [E[f_{ik}(t)]]. \tag{15} \]
The \( m^2 \times m^2 \) covariance matrix of \( f \) has a block diagonal structure since the flows emanating from two different nodes are independent. Therefore, by denoting the \( i^{th} \) row of \( F(t) \) as \( f_{i(\cdot)} \), we write
\[ \Sigma_f = \text{Cov}(f) = \begin{bmatrix} \text{Cov}(f_{i(1)}) & 0 & \cdots & 0 \\ 0 & \text{Cov}(f_{i(2)}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Cov}(f_{m(1)}) \end{bmatrix}. \tag{16} \]
The elements of the \( i^{th} \) \( m \times m \) sub matrix on the diagonal, \([\text{Cov}(f_{i(\cdot)})]\), are given each by
\[ \text{Cov}(f_{i(\cdot)}) = \text{Cov}(f_{ij}; f_{ik}) = \begin{cases} \theta_i p_{ij} (1 - p_{ik}) & \text{if } j = k \\ \theta_i p_{ij} p_{ik} & \text{if } j \neq k \end{cases} \]
Therefore, in the absence of any constraints such as (12) or (13), the probability distribution of \( P \) is approximated by \( \mathcal{N}(\mu_F, \Sigma_F) \).

B. Constrained Gaussian Sampling: Equality Constraints

Given the Gaussian approximation to the multinomial distribution, our objective is to draw Monte Carlo samples from the Gaussian distribution subject to admissibility constraints (12) and (13). In [7] we have shown that in order to satisfy equality constraint (12), we sample from a distribution characterized by
\[ f = A_{oc} v - A^# b \tag{17} \]
where \( A_{oc} \) is the orthogonal complement of the matrix \( A \), and \( A^# \) is its pseudo-inverse. The samples \( v \) are drawn from a lower dimensional Gaussian that is characterized by
\[ \mu_v = A^# (f - A_{oc} v) \tag{18} \]
\[ \Sigma_v = A^# \Sigma f (A_{oc})^T \tag{19} \]
The resulting samples can be shown to satisfy constraint (12).

C. Constrained Gaussian Sampling: Inequality Constraints

To ensure that the Monte Carlo samples satisfy inequality constraint 13, we rely on Gibbs sampling [10], [11]. In this approach, points are sequentially sampled in each dimension by drawing from a truncated multi-variable distributions multiple times. This procedure is detailed as follows:

1) Transform the Gaussian distribution \( \mathcal{N}(\mu_F, \Sigma_F) \) in the problem space to the standard form. The transformed random vector \( z \) is characterized by \( \mathcal{N}(0, I) \). This transformation is accomplished via
\[ z = T^{-1} (f - \mu_f) \]
where \( T = UA^2 \), \( U \) and \( A \) are obtained from an eigen decomposition of the covariance, matrix \( \Sigma_F = UAU^T \).
2) Similarly transform the inequality constraints \( Cf + d \leq 0 \) to equivalent constraints in the \( z \)-space according to \( CTz - C\mu_f \leq d \). This results in an equivalent set of constraints \( Cz \leq d \).
3) Apply Gibb’s sampling by iterating through the dimensions of the search space. This reduces the complicated problem of sampling the inequality constrained Gaussian to a series of univariate truncated Gaussian samples. For each dimension \( k \), the bounds on the univariate truncated Gaussians are given by
\[ z^\text{min} = \max_{C_{ij} > 0} \frac{d_i - \sum_{j \neq k} C_{ij} z_j}{C_{ij}} \quad \text{and} \quad z^\text{max} = \min_{C_{ik} > 0} \frac{d_i - \sum_{j \neq k} C_{ij} z_j}{C_{ij}} \]
4) Transform the samples acquired from the \( z \)-space back to the problem space via
\[ f = Tz + \mu_f. \]

This approach is illustrated in fig. (3). The figure shows the contours of the Gaussian probability distribution in the problem and transformed space. By taking Gibbs’ samples in the transformed space via a series of truncated univariate samples, the constraints are respected immediately. In contrast to rejection sampling, the Gibbs sampling approach provides samples that are drawn from the truncated distribution with full efficiency, though each sample is drawn at a higher cost.

D. Estimating \( P \) from Asynchronous Data

To estimate \( P \) in the asynchronous case, we derive an expression for an equivalent inter-sample probability transition matrix, \( P_{\Delta t} \), as a function of the underlying probability transition matrix \( P \), as well as the statistics of the temporal transition parameters. The derived relationship is reversible, meaning that given an estimate of \( P_{\Delta t} \) we can uncover the underlying \( P \), and vice versa.
Recall that we assumed that the statistics of the temporal state-transitions are governed by a Poisson process (4). Consequently, the number of transitions, \( n_j \), that cell \( j \) undergoes in a time period \( \Delta t \) is given by the Poisson distribution:

\[
\Pr(n_j) = e^{-\lambda \Delta t} \cdot \frac{\left( \lambda \Delta t \right)^{n_j}}{n_j!}
\]

If cell \( j \) assumes state \( i \) at time \( t \), then it occupies state \( k \) at the next transition with probability \( p_{ik} \). Therefore \( i^{th} \) row of \( P \), the vector \( p_{i}(:, \cdot) \), represents the probability distribution that the cell occupies each of the states at the next transition conditioned on the current state being \( i \).

\[
\Pr(s_j(t + \Delta t)|s_j(t) = i, n_j = 1; \lambda) = \left( P^T \right)_i e_i
\]

Similarly, the probability distribution after \( n_i \) transitions, this probability distribution is given by

\[
\Pr(s_j(t + \Delta t)|s_j(t) = i, n_j; \lambda) = \left( P^T \right)^{n_j} e_i
\]

where \( e_i \) is the \( i^{th} \) unit vector. To compute the probability distribution for all possible transitions predicted by the Poisson process, we marginalize over all possible values of \( n_j \):

\[
\Pr(s_j(t + \Delta t)|s_j(t) = i; \lambda) = \sum_{n_j = 0}^{\infty} \Pr(s_j(t + \Delta t)|s_j(t) = i, n_j; \lambda)
\]

\[
= \sum_{n_j = 0}^{\infty} \{ \Pr(n_j|\lambda, \Delta t) \cdot \Pr(s_j(t + \Delta t)|s_j(t) = i, n_j; \lambda) \}
\]

Manipulating this expression leads to

\[
\Pr(s_j(t + \Delta t)|s_j(t) = i; \lambda) = \sum_{n_j = 0}^{\infty} \left\{ e^{-\lambda \Delta t} \cdot \frac{\left( \lambda \Delta t \right)^{n_j}}{n_j!} \cdot \left( P^T \right)^{n_j} e_i \right\} = e^{-\lambda \Delta t} e^{\left( \lambda \Delta t \cdot P^T \right)} \cdot e_i
\]

where the last equality follows from the definition of matrix exponentiation \( e^M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{M^k}{k!} \).

Consequently, an equivalent probability transition matrix that takes into account the asynchronicity of the process is given by

\[
P_{\Delta t} = e^{-\lambda \Delta t} \cdot e^{\left( \lambda \Delta t \cdot P \right)}
\]

The resultant equivalent probability transition matrix, \( P_{\Delta t} \), can therefore be estimated using the tools and techniques developed to estimate the transition probabilities for the asynchronous case. Given \( P_{\Delta t} \), the nominal matrix \( P \) can be recovered according to

\[
P = \frac{1}{\lambda \Delta t} \cdot \log_{M} \left( P_{\Delta t}e^{-\lambda \Delta t} \right)
\]

where \( \log_{M} \) denotes the matrix logarithm.

V. SUMMARY OF THE ALGORITHM

A summary of the algorithm for estimating the probability transition matrix \( P \) from asynchronous, longitudinal observations is as follows:

1) For each observation \( \theta(t) \), approximate the multinomial likelihood function of the flow matrix \( F \) (7) by an equivalent multivariate Gaussian according to equations (15) and (16). Thus \( \Pr(F(t)|P, \theta(t)) \approx N(\mu_F, \Sigma_F) \).

2) Approximate the summation over the set \( F \) by a smaller number of samples

\[
\Pr(\theta(t + 1)|P, \theta(t)) \approx \sum_{F \in \mathcal{F}} N(\mu_F, \Sigma_F)
\]

3) To obtain the samples \( F_{\text{samples}} \):
   a) Compute an equality constrained representation of the Gaussian model using Equations (18) and (19).
   b) Using the procedure of Section (IV-C), obtain inequality constrained Gaussian samples of the sample space.

4) Use the samples \( F_{\text{samples}} \) to compute the posterior distribution according to eq. (14).

5) Adjust the effects of asynchronicity by converting \( P_{\Delta t} \) to \( P \) via equation (22).

VI. SIMULATION RESULTS & DISCUSSION

We simulated the algorithm for a two state process, using the following probability transition matrix as a representation of the real transition parameters:

\[
P_1 = \begin{bmatrix}
0.35 & 0.65 \\
0.80 & 0.20
\end{bmatrix}
\]

In this two-dimensional instance of the problem, all possible probability transition matrices can be succinctly summarized on a two dimensional grid ranging from 0 to 1 on each axis. Figure 4 shows the progression of the algorithm over three time steps as estimates converge to the true value.

While our illustrations focused on 2-state examples for clarity of presentation, it can be extended in a straightforward manner to problems with higher number of states. The computation time for estimating the likelihood of any one candidate transition probability is not a function of the number of states, but rather is primarily a function of the number of samples required to get an accurate summation in the Monte Carlo step. As the number of states increases, however, the number of candidate transition probabilities increases exponentially.

Therefore, the formulation presented in this approach is a generalization of two particular problems addressed prior:

1) The synchronous case is a special one where \( n = 1 \).
2) The case of having full panel data case is a special case when the set \( \mathcal{F} \) admits only one element.
Fig. 3. Illustration of Gibbs sampling scheme to efficiently sample from Gaussian distributions subject to inequality constraints. (a) Gaussian distribution contours shown in problem space. The feasible region satisfying all inequality constraints is shaded in gray. Samples are drawn from the distribution irrespective of the inequality constraints (red asterisks). Only the samples satisfying the inequality constraints are picked (blue circles) and retained. (b) Distribution, data and constraints of subfigure (a) but are shown under the transformation. (c) In the transformed space, samples are drawn from the constrained distribution via Gibbs sampling. All samples satisfy the inequality constraints. (d) Samples are transformed back to problem space, maintaining feasibility.

The algorithm developed here can be directly applied to those two special cases without the need for further development.

It is important to note that the estimation scheme presented here for the computation of the probability of a particular set of parameters is independent of the dimensionality of the problem and also independent of the number of agents considered. Thus, the algorithm scales well with respect to those two parameters, but scales linearly with the number of samples acquired during the Gibbs steps of the algorithm.

REFERENCES