J. Indian Inst. Sci., Mar.–Apr. 2000, **80**, 75–97

Modelling video telephony by Fokker–Planck equation and their analysis

N. SAI SHANKAR* AND A. P. SHIVAPRASAD
Department of Electrical Communication Engineering., Indian Institute of Science, Bangalore 560012, India
email : aps@ece.iisc.ernet.in

Abstract

Video services will form a substantial portion of the traffic carried by emerging broadband integrated networks. The video streams are compressed and sent to the network to conserve bandwidth and hence require a good model to represent them in an asynchronous transfer mode (ATM) network. In this paper, we model the video sources by means of a diffusion processes with mean μ, variance σ^2 and co-variance $C(t)$. We consider video teleconferencing where the changes in the successive frames are not abrupt but are more or less uniform. The matching parameters for the mean, variance and covariance are taken from discrete-time Markov chain (DTMC) found in Blondia and Casals (*Performance Evaluation*, 1992, **16**, 5–20). The video telephony session is modelled according to a truncated diffusion process and we calculate the steady-state behaviour for such a process. We consider two truncations, namely, reflecting barrier and absorbing barrier and state their intuitions in video telephony.

Keywords: Diffusion process, traffic models, VBR sources.

1. Introduction

The asynchronous transfer mode (ATM) is considered by the International Telecommunication Union-Technology (ITU-T) Standards Committee as the transport method for the broadband integrated services digital network (B-ISDN). In this network, digital video communication(video phone, video conferencing, television distribution, etc.) is expected to be a major class of service. ATM offers the capacity to support variable bit rate (VBR) connections. These input sources share a link of constant capacity d, which is less than the sum of individual bit rates, achieving a significant multiplexing gain. In particular, information loss and delay are the most important parameters that determine the quality of service. This paper derives efficient diffusion models for the VBR traffic representing the video traffic offered in an ATM network.

Traffic is the driving force of communication systems, and traffic models are of crucial importance for assessing network's performance. In practice, stochastic models of traffic systems are relevant to network traffic engineering and performance analysis, to the extent that they are able to predict system performance measures to a reasonable degree of accuracy. The fundamental systems, of which traffic is a major ingredient, are queueing systems.

Fluid flow models have been used extensively in modeling high-speed communication networks.[1, 2] The fluid represents packets or bits of information. The fluid flow model de-

*Present address: Philips Research Laboratories, WY 21, Prof. Holstlaan 4, 5656 AA, Eindhoven, The Netherlands
email : sai.shankar@philips.com

scribes the stochastic behaviour of the fluid flow in the buffer. The fluid arrival process is modelled by Markov-modulated process which is described by a continuous time Markov chain (CTMC).[3] In spite of their success as modelling tools in ATM networks, the fluid flow models have one drawback, that is they are second-order models. Hence, they work well when the system characteristics are adequately described by first- and second-order moments. This is the reason why they are very attractive in modelling D/D/1 queue in random environment. Consider a queue with a buffer. When the number of customers in the queue is very high, each discrete step's increase in the queue length, because of arrivals and decrease in queue length because of departures, is very small when compared to the total queue length. In such cases it is reasonable to approximate the discrete flow of customers by a continuous flow like a fluid. Any such results originate from fluid mechanics, in particular the diffusion for an ideal gas. Hence we call them as diffusion approximations. These techniques are also useful is studying the transient behaviour of the queueing system. In ATM networks, where the packets are of constant size and sources behave in an "on-off" fashion, the first-order models cannot account for the variablity of the arrival rate of the fluid. This is discussed in Karandikar and Kulkarni.[4] It is this limitation that has prompted us to investigate the second-order fluid flow models to model ATM networks. In the case of VBR models, there is a positive serial dependence between successive packet arrivals, and this dependency is a major cause of congestion in multiplexer queue.

In this paper, we analyze the traffic patterns of the stochastic fluid flow system shown in Fig. 1. The system consists of a buffer of finite size B, with the server having a constant rate of output d, which receives input from heterogeneous set of traffic sources. Our primary motivation for considering this system stems from its use in modelling statistical multiplexing of several sources at an ATM switching element in a B-ISDN. The high rates of transfer in such networks make conventional queueing models, which treat individual cell as customer, unwieldy. The input flow characterized by diffusion process, accurately captures the bursty nature of data flow in such networks. The idea of approximating a discrete state process by a diffusion process with continuous path was discussed in detail in Feller.[5] The procedure of using a diffusion process to study a queueing system, whether it be a continuous time system or a discrete-time system can be useful because mathematical methods associated with continuum very often lend themselves more easily to analytical treatment than those associated with discrete coordinate axes. Further references on diffusion model systems are available elsewhere.[1, 6–9]

1.1. *Motivation for considering the diffusion process for moving pictures expert group (MPEG) streams*

Traffic characterization and source modelling of VBR-coded video are active research areas as VBR video would be a major player in the future of B-ISDN. We consider the MPEG-2 VBR video source. It was pointed out by Lou *et al.*[10] that the frame sequences are normally distributed when we make the following normalizations. The MPEG video can be grouped into three types, I, P and B frames. These frames are different from each other and have different sizes. An MPEG-2 video stream contains the mixture of all the three frames with periodical frame structure. As a typical example we consider the 'Mobl' sequence shown in Fig. 2. The normalization procedure is a linear process and is explained below.

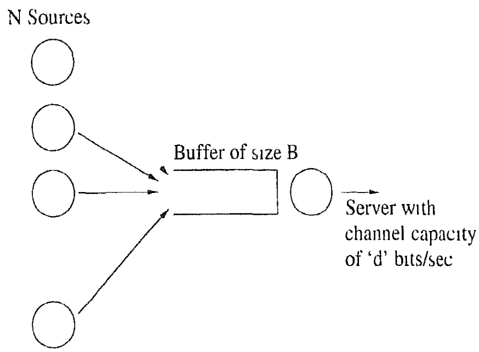

FIG 1. Schematic of an ATM switch

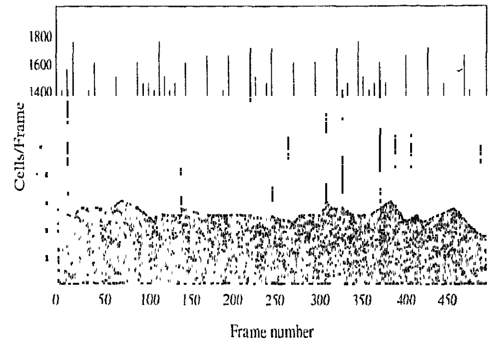

FIG 2 Mobl sequence

Let M_I, M_P and M_B denote the mean size of I, P and B frames, respectively. Let $f(i)$ denote the ith frame appearing in the MPEG stream. Let $\hat{f}(i)$ denote the ith frame appearing in the MPEG stream after normalization. Define

$$S_{IB} = \frac{M_I}{M_B}, S_{PB} = \frac{M_P}{M_B}. \tag{1}$$

The following linear transformation is applied to each frame in the sequence

$$\hat{f}(i) = \frac{f(i)}{S_{IB}}$$

if the ith frame is a B frame,

$$\hat{f}(i) = \frac{f(i)}{S_{PB}} \tag{2}$$

if the ith frame is a P frame, and

$$\hat{f}(i) = f(i)$$

if the ith frame is an I frame. After normalization, the number of cells per frame in the whole sequence is at a comparable level. It is assumed that the data series of the frame size after the normalization can be described by the normal distribution as shown in Fig. 3. The histogram of the 'Mobl' sequence is plotted and the distribution is found to be normally distributed. To test whether the marginal distribution of the frame size is indeed a normal distribution, a Q-Q plot which plots the quantiles of data vs the quantiles of the fitted distribution is also used (Fig. 4). The fit is fairly good except for a few points. This test was carried out for most of the sequences and the results are similar. So the conclusion is that the MPEG sequences can be described by the normal distribution. The linear transformation does not change the distribution type of the series if it follows a normal distribution. Simonian[6] had indicated that experimental measures on animated picture transmission in packet mode have shown that the input process

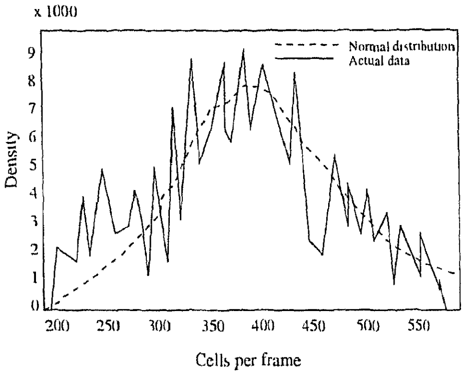

FIG 3 Histogram of cells/frame vs normal distribution

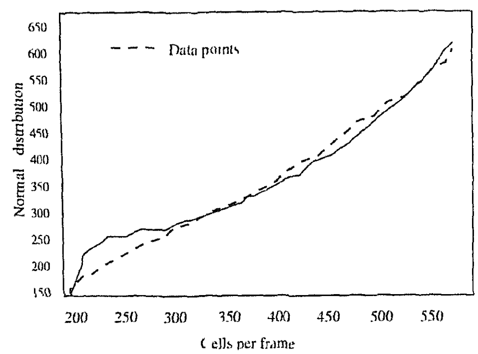

FIG. 4. Q-Q plot of cells per frame vs normal distribution.

can be modelled by a continuous gaussian stationary process with constant mean μ and the decreasing exponential correlation function given by

$$E[(I_{t+\delta t}-\mu)\cdot(I_t-\mu)]\ \Sigma^2\cdot e^{-\eta|\tau|} \tag{3}$$

where I_t is the instantaneous arrival process, μ the mean of the arrival process, η the rate parameter, τ the time constant and Σ is a constant. Simonian[6] also points out that the theoretical parameter σ and the measured parameters Σ and η are related by

$$\sigma^2 = 2\eta\Sigma^2 \tag{4}$$

where $\eta = \alpha + \beta$. Here α and β represent the probability that 'on-off' source is 'silent' or 'active', respectively, in the next time period. Also, let $\tau = \eta t$. So the queue length process is scaled as $\hat{Q}_\tau = \eta Qt/\Sigma$. So, whenever we use the fluid content in the buffer, we use $\hat{Q}_\tau$, instead of Q_t.

2. Diffusion approximation for VBR sources

In this section, we propose analytical models for variable bit-rate video sources and their superpositions. These models are based on the earlier work of Maglaris *et al.*[3] Maglaris *et al.*[3] and Blondia *et al.*,[8] approximate video sources and their superpositions by means of a discrete-state continuous-time Markov process. In this paper, we model video sources by diffusion process which is very useful for fluid approximation approach in statistical multiplexing. We consider sources with uniform activity level, *i.e.*, where no sudden changes in bit rate occur. The typical example is the video phone where there are no abrupt changes in the scene, except for the minor movements of head and shoulder.

2.1. *Video sources with uniform activity level*

Consider a video scene where the changes are uniform. The scene is composed of M basic information blocks which are either 'on' or 'off' depending on the scene characteristics. Mathematically speaking, we have M identical independent 'on-off' sources each generating A bits/s when in 'on' state. The continuous time version of the 'on-off' source is shown in Fig. 5. We

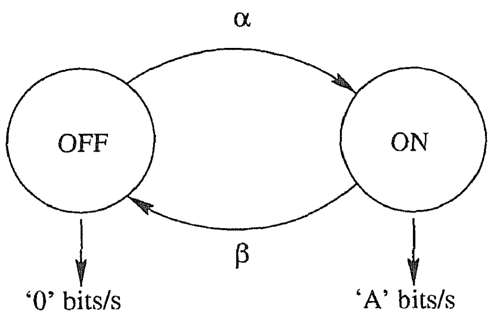

FIG 5 "on-off" source model

consider the discrete time case, which is represented by a two-state Markov chain similar to the continuous time case. Here we have probabilities instead of rates. Also there are probabilities for the remaining in the same state. They are geometrically distributed which is the discrete-time version of the exponential distribution. We consider the versatile point process called discrete-batch Markovian arrival process (D-BMAP), which is the discrete time analogue of the batch Markovian arrival processes introduced by Lucantoni.[11] It has the advantage of modelling large classes of VBR sources.

Some good examples of the slow change scenes available in practice are the bike, the float and the monkey sequences. Here the successive scene frames do not change very rapidly, but do so very slowly. The sequences are shown in Fig. 6. The values of mean, variance and co-variance are derived by calculating the steady state mean, variance and co-variance of the Markov chain as shown in Fig. 9. Figure 9 is a CTMC but we consider it to be a discrete time Markov chain (DTMC). The superposition of M sources constitutes a video telephone session. These 'on-off' sources are multiplexed on a link of capacity d bits/s. The cell-arrival rate is quantized into discrete level of A cells/s, i.e. when the state of the Markov chain is in state i, $(0 \leq i \leq M)$, the source emission rate is iA bits/s. We assume $M+1$ possible levels $\{0, A, 2A, \ldots\ldots, M A\}$ and the transitions take place between neighbouring quantization levels only. Furthermore, we assume that the transition occurs with geometrical transition rates. So the cell stream arriving at the multiplexer buffer is approximated by a superposition of M identical independent 'on-off' sources, called mini sources, each generating cells at a constant rate while being in geometrically distributed active period. The resultant Markov chain is equivalent to sum of M identical two-state mini sources each moving back and forth geometrically between

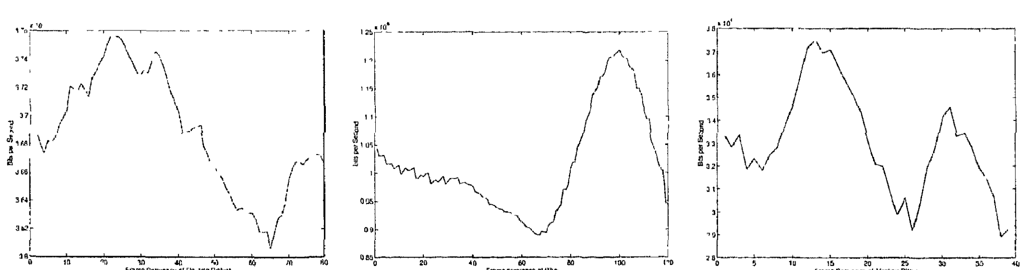

FIG. 6 Frame sequences of monkey, bicycle and floating pictures.

an 'off' and an 'on' state in which A bits/pixel are offered to the access buffer. The composite rate, in the units of bits/pixel, is labelled as I_t. Note that the rate will be changing in frame intervals (30 frames/s) rather than continuously. This is also the multiplexed rate into the buffer. The other way to look into the diffusion approximation is to treat the individual pixels as independent 'on-off' sources having the rates of transitions as that of a two-state DTMC or CTMC. The transition rate matrix as given by Blondia and Casals[8]

$$P=\begin{bmatrix} 1-M\alpha & M\alpha & 0 & \cdot & \cdot & 0 \\ \beta & 1-\beta-(M-1)\alpha & (M-1)\alpha & 0 & \cdot & 0 \\ 0 & 2\beta & 1-2\beta-(M-2)\alpha & (M-2)\alpha & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & \cdot & M\beta & 1-M\beta \end{bmatrix} \tag{5}$$

This captures the model more accurately. As the number of pixels is very large, the resultant equations arising from the usual Markov chain approximation is very hard to solve.

The continuous time analogue of the 'on-off' source is shown in Fig. 5. Now we construct a diffusion model as a limit for the the model given in Blondia and Casals.[8] Consider the DTMC and its associated transition probability matrix as given in eqn (5).

$$\begin{aligned} P_{i,i+1} &= (M-i)\alpha \\ P_{i,i} &= 1-i\beta-(M-i)\alpha \\ P_{i,i-1} &= i\beta. \end{aligned} \tag{6}$$

Given that the Markov chain is in state i, it can go to $i+1$ or $i-1$ state or remain in the same state with the above probabilities. The infinitesimal mean rate of change from a state i is given by

$$[P_{i,i+1} - P_{i,i-1}]\Delta = -\mu(i\Delta)\Delta t = [(M-i)\alpha - i\beta]\Delta. \tag{7}$$

The infinitesimal variance of the change from one state to another is given by

$$[P_{i,i+1} + P_{i,i-1} - (P_{i,i+1} + P_{i,i-1})^2]\Delta^2 = \sigma^2(i\Delta)\Delta t. \tag{7}$$

Using the standard technique of converting a birth-death process into a diffusion process as indicated in Cox and Miller,[12] and appendix A.1, we get

$$\frac{\partial f(t;x,y)}{\partial t} = \mu(x)\frac{\partial f(t;x,y)}{\partial x} + \frac{\sigma^2(x)}{2}\frac{\partial^2 f(t;x,y)}{\partial x^2}. t > 0.0 < x, y < M \tag{9}$$

where $\mu(x)$ is replaced by the steady-state mean fluid rate of the sources given by $\mu = \frac{MA\alpha}{\alpha+\beta}$ and similarly $\sigma^2(x)$ is replaced by the steady-state variance $\sigma^2 = \frac{MA^2\alpha\beta}{(\alpha+\beta)^2}$, since we use moment-matching technique to analyze the performance of our system with the system given in Anick *et al.*[1] and Simonian.[6] $f(t; x, y)$ denotes the conditional probability density of the arrival process with instantaneous arrival rate of x at time t given that arrival rate is y at time 0. Though the above equation is governed by the infinitesimal mean and variance which is dependant on the state of the Markov chain and is given by eqns (7) and (8), respectively, we have plugged the steady state mean and variance values into the diffusion equation and analyze the performance of such systems. The boundary conditions ensure that eqn (9) is 0 at 0 and M. The matching parameters for μ, $\sigma\{2\}$ and $C(t)$ are found in the same way as in Maglaris *et al.*[3] The Fokker–Planck equation derived above can also be treated as a stochastic differential equation of the form

$$dI_t = \mu(I_t)dt + \sigma(I_t)dW_t.$$

the above equation, W_t, μ and σ^2 are the standard Wiener process, steady-state mean of the process I_t and steady state variance of the process I_t, respectively.

3. Calculation of the transition probabilities

The diffusion equation modelled above is the result of applying limits on a finite birth–death process whose transition rates are given in eqn (5). We now proceed to derive the steady-state distribution of the instantaneous arrival rate process I_t described by the stochastic differential equation given by eqn (10) under the conditions that $\sigma > 0$. In order to discuss the boundary condition and the stationary distribution, we shall approximate the process I_t by a sequence of diffusion processes with smooth drift and diffusion coefficients. Let $\mu^n(I)$ and $n(I)$ be monotone and sufficiently smooth functions such that $\mu^n(I) = \mu(I)$ and $\sigma^n(I) = \sigma(I)$ for I outside the interval $(-n^{-1}, n^{-1})$. Then the stochastic differential equation

$$dI_t^n = \mu^n\left(I_t^n\right)dt + \sigma\left(I_t^n\right)dW_t \tag{11}$$

has a unique solution I_t^n. This is explained by Wong.[13] Since $\mu^n \to \mu$ and $\sigma^n \to \sigma$ as $n \to\infty$ except for $I = 0$, the sequence of stochastic processes $\left\{I_t^n\right\}$ converges to a limiting process I_t. It is clear that I_t satisfies the original stochastic differential equation [eqn (10)]. So we shall take the limit I_t as the solution of eqn (10). If the dependence of the mean and variance on x is considered, then we would end up with Ornstein–Uhlenbeck process, whose steady-state density is gaussian and is given by Kobayashi and Ren,[7]

$$f(y) = \lim_{t\to\infty} f(y,t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-y^*)^2}{2\sigma^2}} \tag{12}$$

where $y^* = \frac{N\alpha}{\alpha+\beta}$ and $\sigma^2 = \frac{N\alpha\beta}{(\alpha+\beta)^2}$. We consider the diffusion model by plugging the steady-state mean and state variance of the original Markov chain into the mean and variance of the diffusion process and then analyzing the behaviour of the diffusion process.

3.1. *Case 1: Both the boundary points are reflecting*

The above diffusion equation has boundaries at 0 and M. We can assume that the boundaries 0 and M are either reflecting or absorbing. If we assume the former, then the process does not stay at the boundaries and is reflected as soon as it reaches any one of the two boundaries. If

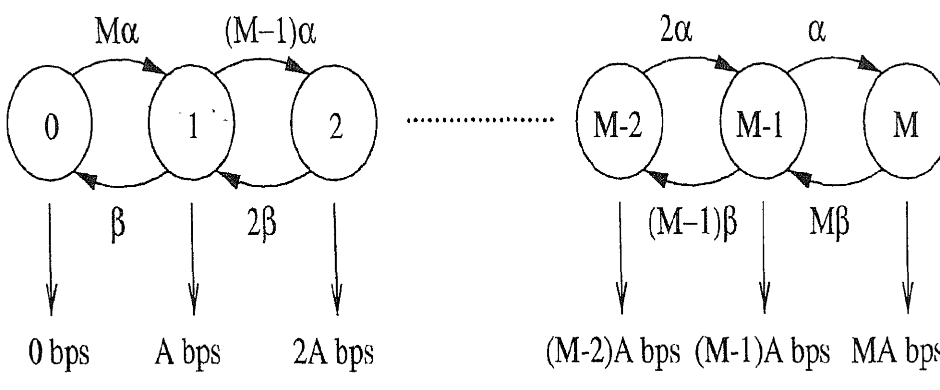

FIG. 7. Model for combined M 'on-off' sources.

the assumption is the latter, then we have to device ways of making the process to come out of the absorbing boundaries. Gelenbe[14] derives solutions with such equations. We can consider the boundaries M and 0 to be reflecting boundaries. The intuition in choosing a reflecting boundary at point M stems from the fact that the information sent at this point corresponds to the starting or refreshing of a new scene, which requires transmission of the entire data from the encoder to the network. The boundary conditions are given by

$$\frac{\partial f(x;t)}{\partial t} = 0 \text{ at } x = 0 \quad (t > 0);$$
$$\frac{\partial f(x;t)}{\partial x} = 0 \text{ at } x = M \quad (t > 0). \tag{13}$$

The solution of the conditional probability density of the unrestricted process starting at x_0 at time 0 and reaching x at time t is given by

$$f_1(x, t / x_0, 0) = \frac{1}{\sigma\sqrt{2\pi t}} e^{-\frac{(x - x_0 - \mu t)^2}{2\sigma^2 t}}. \tag{14}$$

Using the method of images as in Sommerfeld[15] and Cox and Miller[12] the conditional transition probability density for the diffusion process with the reflecting boundary [0, M] is given by

$$f(x, t / x_0, t_0) = \sum_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi t\sigma^2}} \left[e^{-\frac{(2m + x - x_0 - \mu t)^2}{2\sigma^2 t}} + e^{-\frac{(2m + x + x_0 - \mu t)^2}{2\sigma^2 t}} \right] \tag{15}$$

for x, $x_0 \in [0, M]$. The conditional pdf with intital conditions $x = 0$ and $t = 0.01$ is plotted in Fig. 8(a). The plot is also verified using the Mathematica package.

3.3. *Case 2. Only 0 as reflecting barrier*

Now consider the case when the number of 'on-off' sources, $M \to \infty$, $\alpha \to 0$ and $\beta \to 0$ in such a way that $M\alpha = a$ and $M\beta = b$ where a and b are constants. In this case we get a diffusion

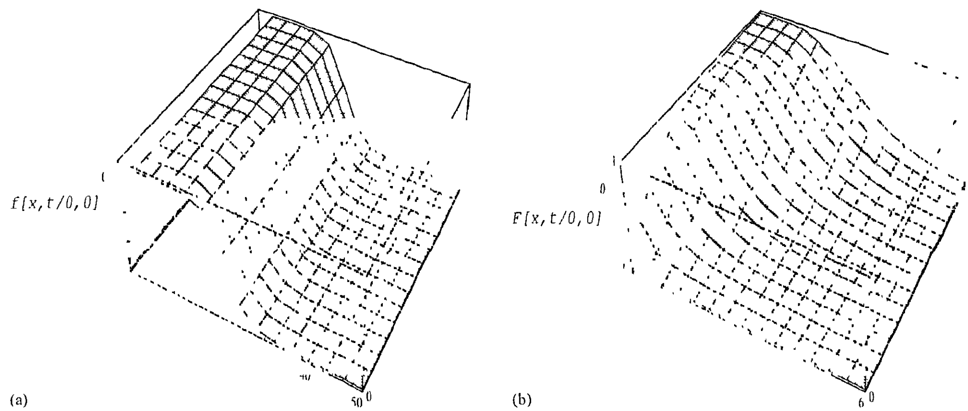

FIG 8(a). Conditional probability density function $f(x, t/0, 0)$ and (b) Conditional probability distribution function $F(x, t/0, 0^{+})$. The mathematica package was not able to plot the spike at $(x, t) = 0, 0$

approximation as in the previous case but with the reflecting barrier at M being removed. The conditional transitional probability density of the arrival rate process $f(x, t/x_0, 0)$ as given by Cox and Miller[12] is

$$f(x,t/x_0,0)=\frac{1}{\sqrt{2\pi\sigma^2 t}}\left[e^{-\frac{(x-x_0-\mu t)^2}{2\sigma^2 t}}+e^{-\frac{4x_0\mu t-(x-x_0-\mu t)^2}{2\sigma^2 t}}+\frac{2\mu}{\sigma^2}e^{\frac{2\mu x}{\sigma^2}}\{1-\phi\left(\frac{x-x_0+\mu t}{\sigma\sqrt{t}}\right)\right] \quad (16)$$

for $x, x_0 \in [0. \infty)$. (See Appendix A2 for the derivation of eqn (16).) Now let us calculate the moments of the conditional probability density function. From eqn (16), it can be shown that the conditional probability distribution function is given by

$$F(x,t/x_0,0)=P(X(t)\le x/X(0)=x_9)=\frac{1}{2}\left\{1+\mathrm{Erf}\left[\frac{x-x_0-\mu t}{\sigma\sqrt{2t}}\right]-e^{\frac{2\mu x}{\sigma^2}}\left(1+\mathrm{Erf}\left[\frac{x-x_0-\mu t}{\sigma\sqrt{2t}}\right]\right)\right\} \quad (17)$$

where $\mu = (a-b)A$. The probability distribution function $F(x, t/0, 0)$ is plotted in Fig. 8(b). The steady-state pdf $I(x)$ and the asymptotic distribution function are found to be

$$I(x)=\lim_{t\to\infty} f(x,t/x_0,0)=\frac{2\mu}{\sigma^2}e^{-\frac{2\mu x}{\sigma^2}} \quad (18)$$

and

$$F_X(x) = P(X(\infty) < x) = 1 - e^{-\frac{2\mu x}{\sigma^2}}. \quad (19)$$

The above is an excellent model for the video telephony where the changes in the scenes are very slow. This is evident from the plots of density and distribution functions. The steady-state behaviour of the input is exponentially distributed.

3.3. *Steady-state behaviour of the input rate*

Clearly the distribution of I_t at t converges to the stationary distribution as $t \to \infty$. Let us denote $f(x/x_0)$ as the steady-state conditional probability density function. Then as $t \to \infty, \frac{\partial f(x/x_0)}{\partial t} \to 0$. Hence the Fokker–Planck eqn (9) becomes

$$\mu(x) f(x/x_0) + \frac{\sigma^2(x)}{2} \frac{d}{dx} f(x/x_0) = 0. \tag{20}$$

Solving the above equation we have

$$f(x/x_0) = \frac{C}{\sigma^2(x)} e^{\int_0^x -2\frac{\mu(y)}{\sigma^2(y)} dy} \tag{21}$$

where $\mu(x)$ and $\sigma^2(x)$ are replaced by the steady-state values μ and σ^2, respectively, as given in section 2.1. The constant C is determined from the total probability condition, namely, $\int_0^\infty f(x/x_0) = 1$. So C is found to be

$$C^{-1} = \frac{1}{\sigma^2(x)} \int_0^\infty e^{\int_0^x -2\frac{\mu(y)}{\sigma^2(y)} dy} dx. \tag{22}$$

Now $f(x/x_0)$ is given by

$$f(x/x_0) = \frac{C}{\sigma^2} e^{-\frac{2\mu}{\sigma^2} x}. \tag{23}$$

This expression agrees with the expression derived in the previous section. Now we can also determine the conditional probability of the input rate exceeding the output rate given that the input rate is less than the output rate, i.e. $f(x > d/x_0)$.

3.4. *Case of 0 as an absorbing boundary*

The case of having absorbing boundaries at point 0 is tantamount to treating the video connection as closed. This is a fairly good assumption for the case where the connection ceases after some finite time. To derive the solution for the condition with the absorbing boundary at the origin, we consider the case of two absorbing boundaries 0 and d and then extend it to the case of a single absorbing boundary. This implies that the following boundary condition holds.

$$\lim_{x \to 0} f(x/x_0) = 0 \quad \text{and} \quad \lim_{x \to d} f(x/x_0) = 0. \tag{24}$$

We use the method of the Fourier series to solve the above equation and find that the functions

$$\psi_m(x) = b_m e^{\frac{\mu x}{\sigma^2}} \sin\left(\frac{m\pi x}{d}\right) \tag{25}$$

$$\alpha = \alpha_m = -\frac{m^2\pi^2\sigma^2}{2d^2} - \frac{\mu^2}{2\sigma^2}. \tag{26}$$

The absorbing boundaries are called Dritchlet boundaries and the steady-state probability is given by

$$\pi(x) = e^{\frac{2\mu x}{\sigma^2}} \text{ for } (0 \le x \le d). \tag{27}$$

The normalizing constants are

$$b_m = \sqrt{\frac{2}{d}}(m = 1,2,\ldots). \tag{28}$$

The eigenfunction expansion of f is given by

$$f(x / x_0) = \sum_{m=1}^{\infty} e^{\alpha_m t} \psi_m(x) \psi_m(x_0) \pi(x_0) \tag{29}$$

$$= \frac{2}{d} e^{\frac{\mu(x_0 - x)}{\sigma^2}} e^{-\frac{\mu^2 t}{2\sigma^2}} \sum_{m=1}^{\infty} e^{-\frac{m^2\pi^2\sigma^2 t^2}{2d^2}} \times \sin\left(\frac{m\pi x}{d}\right) \sin\left(\frac{m\pi x_0}{d}\right).$$

Now if we take the limit of $d \to \infty$ we get the diffusion process with absorbing barrier at 0.

$$f(x / x_0) = 2e^{\frac{\mu(x_0 - x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}} \int_0^{\infty} e^{-\frac{\pi^2\sigma^2 t u^2}{2}} \sin(\pi x u) \sin(\pi x_0 u) du \tag{31}$$

$$= \frac{2}{\pi} e^{\frac{\mu(x_0 - x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}} \int_0^{\infty} e^{-\frac{\sigma^2 t \zeta^2}{2}} \sin(\pi x \zeta) \sin(\pi x_0 \zeta) d\zeta \tag{32}$$

$$= \frac{1}{\pi} e^{\frac{\mu(x_0 - x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}} \int_0^{\infty} e^{-\frac{\sigma^2 t \zeta^2}{2}} \left[\cos(\zeta(x - x_0)) - \cos(\zeta(x + x0))\right] d\zeta \tag{33}$$

$$= e^{\frac{\mu(x_0 - x)}{\sigma^2} - \frac{\mu^2 t}{2\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2 t}} \left[e^{-\frac{(x - x_0)^2}{2\sigma^2 t}} - e^{-\frac{(x + x_0)^2}{2\sigma^2 t}} \right] \tag{34}$$

3.5. *Calculation of loss probability in case of zero buffer for single source*

From Schwartz,[16] we know that the loss probability for the model mentioned in Fig. 7., P_L, is given by

$$P_L = \sum_{i=L+1}^{M} (i - d)\pi_i / m \tag{35}$$

where m is the mean number of the sources that are in the 'on' state, $J = \{L + 1, L + 2, \ldots\ldots, M\}$ represents the overload region, i.e. it represents those states in the Markov Chain in which the arrival rate is greater than the service rate d. Let π_i represent the steady-state probability that the Markov Chain is in state i. This is shown in Fig. 9. Another related measure, ε as indicated in Schwartz[16] is given by

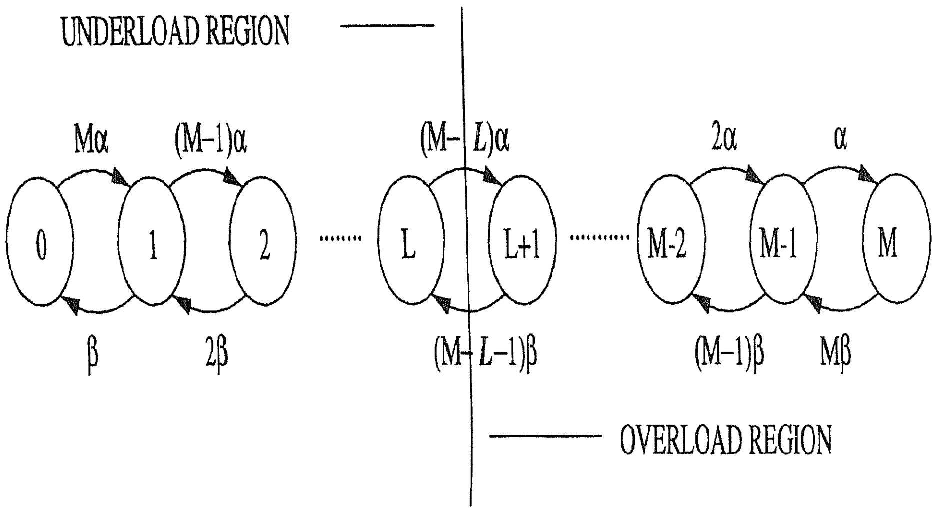

FIG. 9. Model showing the under and overload regions

$$\varepsilon = \sum_{i=L+1}^{M} \pi_i . \tag{36}$$

This is simply the probability of being in the overload region. Now in the diffusion model in which the input is unrestricted, we have

$$P_L = \frac{1}{\mu} \int_{L+1}^{\infty} \frac{e^{-\frac{(x-x_0-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi t}\,\sigma} (x-d)\,dx \text{ and} \tag{37}$$

$$\varepsilon = \int_{L+1}^{\infty} \frac{e^{\frac{-(x-x_0-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi t}\,\sigma} dx. \tag{38}$$

Rewriting $(x-d)$ as $(x-x_0-\mu t)-(d-x_0-\mu t)$ and substituting in the above equation we have

$$P_L = \frac{1}{\mu} \int_{d}^{\infty} \frac{(x-x_0-\mu t) e^{-\frac{(x-x_0-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi t}\,\sigma} dx - \frac{(d-x_0-\mu t)}{\mu} \varepsilon \tag{39}$$

$$= \frac{\sigma t}{\mu} \frac{e^{\frac{-(d-x_0-\mu t)^2}{2\sigma^2 t}}}{\sqrt{2\pi t}} - \frac{(d-x_0-\mu t)\varepsilon}{\mu}. \tag{40}$$

In the case of restriction with absorbing boundaries, the cell loss probability is given by

$$P_L = \frac{1}{\mu}\int_d^\infty e^{\frac{\mu(x_0-x)}{\sigma^2}-\frac{\mu^2 t}{2\sigma^2}} \frac{1}{\sqrt{2\pi\sigma^2 t}}\left[e^{-\frac{(x-x_0)^2}{2\sigma^2 t}} - e^{-\frac{(x+x_0)^2}{2\sigma^2 t}}\right]dx. \tag{41}$$

Similarly, we can plug the value of the process with reflected boundaries in eqn (35) and get the cell-loss probability.

3.6. *Superposition of N such sources*

We consider N-such video telephony connections each represented by the M-state Markov chain as said in the above section, with each session having its own mean μi, variance σ_i^2 and covariance $C_i(t)$. Also, we assume that each session is independent of other sessions and get the diffusion equation of the following form

$$\frac{\partial p(t;x,y)}{\partial t} = \sum_{i=1}^{N} \mu_i(x)\frac{\partial p_i(t;x,y)}{\partial x} + \sum_{i=1}^{N} \frac{\partial_i^2(x)}{2}\frac{\partial^2 p_i(t;x,y)}{\partial x^2}.t > 0.0 < x, y < \infty. \tag{42}$$

Theorem 3.1. The limiting densities p_j of the above processes satisfy the following equation

$$0 = \sum_{i=1}^{N} \mu_i(x)\frac{\partial p_i(x)}{\partial x} + \sum_{i=1}^{N} \frac{\partial_i^2(x)}{2}\frac{\partial^2 p_i(x)}{\partial x^2}.t > 0.0 < x, y < \infty \tag{43}$$

along with the boundary conditions

$$0 = \sum_{i=1}^{N} \mu_i(x)\frac{\partial p_i(0)}{\partial x} + \sum_{i=1}^{N} \frac{\partial_i^2(x)}{2}\frac{\partial^2 p_i(0)}{\partial x^2}.t > 0, \tag{44}$$

and

$$0 = \sum_{i=1}^{N} \mu_i(x)\frac{\partial p_i(M)}{\partial x} + \sum_{i=1}^{N} \frac{\partial_i^2(x)}{2}\frac{\partial^2 p_i(M)}{\partial x^2}.t > 0. \tag{45}$$

The above boundary conditions are got by considering reflecting boundaries at 0 and M.

Proof: Introduce the matrix notation as follows

$$A = \frac{1}{2}\begin{bmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \sigma_N^2 \end{bmatrix}, B = \begin{bmatrix} \mu_1 & & & & \\ & \mu_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \mu_N \end{bmatrix} \text{ and } p = [p_1, p_2, \ldots\ldots, p_N].$$

This can be written as

$$A\ddot{p}(x) + B\dot{p}(x) = 0 \text{ and}$$
$$A\ddot{p}(0) + B\dot{p}(0) = 0 \text{ and}$$

$$A\ddot{p}(M) + B\dot{p}(M) = 0 \tag{46}$$

where $\ddot{p}(x)$ and $\dot{p}(x)$ are first and second derivatives of $p(x)$ respectively. From the general theory of linear differential equations we know that the solution of the eqn (46) is a linear combination of the functions of the type

$$p(x) = e^{\lambda x}\Phi$$

where λ is a scalar and $\Phi = [\Phi_1, \Phi_2, \ldots\ldots \Phi_N]$. Substituting eqn (47) in eqn (46) we get

$$\Phi[\lambda^2 A + \lambda B] = 0 \tag{48}$$

Thus, λ should be a solution to

$$\det[\lambda^2 A + \lambda B] = 0 \tag{49}$$

and Φ satisfies eqn (48). Thus the superposition process, in steady state, is also a truncated gaussian process with the mean given by $\mu = \sum_{i=1}^{N} \mu_i$ and variance by $\sigma^2 = \sum_{i=1}^{N} \sigma_i^2$. The solution to the superposition of N identical sources is also the same as that of the individual process. So the expressions derived for the single source case is applicable for these sources except the mean and variance of superposition is the sum of mean and variance of individual sources.

3.7. *Calculation of loss probabilities for the zero-buffer case when there are N sources*

As mentioned above, let there be N such sources which emit the fluid to the output port which has no buffer. The ith source emits fluid according to the diffusion process with the mean rate $\mu_i(x)$, variance $\sigma_i^2(x)$ and autocovariance $C_i(x)$. For the sake of simplicity, assume that all the sources are identical and have the same mean, variance and autocovariance. The loss-probability analysis is similar to that of the single source case. We make an assumption that the resultant process is normally distributed and apply the theory of large deviations to calculate the cell-loss probability. Now it is easier to find the probability that the input rate exceeds the output rate. As per the assumption we have N identical independently distributed untruncated normal random variables whose generating function is given by

$$(\theta) = \frac{1}{2}\int e^{\theta y} e^{-\frac{(y-\mu t)^2}{2t\sigma^2}} dy = e^{\mu\theta t + \frac{t\sigma^2\theta^2}{2}}. \tag{50}$$

Now proceeding as given in Shwartz and Weiss[17] we get $l(d) = \sup_\theta\left(\theta d - \left[\mu\theta t + \frac{1}{2}\sigma^2\theta^2 t\right]\right) = \frac{(d-\mu t)^2}{2t\sigma^2}$. Thus Chernoff's Theorem states that, for any $d > 0$,

$$P\left(x_1 + x_2 + \ldots + x_N \geq Nd\right) \approx e^{-\frac{1}{2}N\frac{(d-\mu)^2}{\sigma^2}}. \tag{51}$$

But we need the probability that the N sources, transmitting simultaneously, exceeds the service rate Nd. Here it is assumed that the capacity of the channel in Nd instead of d. In this case we can perform a direct calculation: $x_1 + \ldots\ldots x_N$ is a normal random variable distributed as $\sqrt{N}x_1$, so

$$p\left(x_1 + x_2 + \ldots + x_N \geq Nd\right) = P\left(x_1 \geq \sqrt{N}d\right) = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{N}d}^{\infty} e^{-\frac{1}{2}t^2}\, dt. \tag{52}$$

Using the estimate of this integral as in McKean,[18]

$$\frac{1}{y + y^{-1}} e^{-\frac{1}{2}y^2} \leq \int_y^{\infty} e^{-\frac{1}{2}t^2}\, dt \leq \frac{1}{y} e^{-\frac{1}{2}y^2} \tag{53}$$

we obtain

$$P\left(x_1 + x_2 + \ldots + x_N \geq Nd\right) \approx \frac{1}{\sqrt{2t\pi N}\, d} e^{-N\frac{(d-\mu t)^2}{t\sigma^2}} \tag{54}$$

which is in agreement with the exponential order of the large deviation estimates. The fact the $\frac{1}{\sqrt{N}}$ appears is also generic.

4. Buffer behaviour

4.1. *A diffusion approximation to the fluid model*

In order to get the exact solution for the steady-state distribution of the fluid model with diffusion process as input, one has to deal with a two-dimensional stochastic process $(Q_t, I_t)_t \geq 0$ on $\Re[0, \infty) \times \Re[0, \infty)$. It is very difficult to solve such case. So we make an approximation of the queue length process Q_t by a diffusion process X_t on $\Re(-\infty, \infty)$. The approximating process X_t has no reflecting barrier at 0. So when the process X_t is in the range of $(\infty, 0]$, we treat the original queue length process Q_t as 0. Define an one-dimensional diffusion process X_t by a stochastic differential equation in the form

$$dX_t = \mu(X_t)dt + \sigma(X_t)\, dW_t \tag{55}$$

where W_t is the standard Wiener process and the coeffecient of drift is $\mu(x)$ and coeffecient of diffusion is $\sigma(x)$. The above stochastic differential equation is interpreted as

$$X_t = X_0 + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s \tag{56}$$

where the second integral is Ito's integral.

Consider the model as shown in Fig.1. We consider a fluid buffer model with stochastic instantaneous input rate I_t, constant output rate d and buffer of size B. Here I_t is a diffusion process such that the $\int_0^t I_t d\tau = A_t$, the cumulant input has continuous sample paths. The process A_t represents integrated Wiener process. So the rate of change of A_t follows a diffusion process. When the buffer is not empty the rate of change in the buffer content Q_t, depends on the difference $I_t - d$; and when the buffer is empty, it remains zero until the instantaneous arrival rate I_t exceeds d. So the buffer content Q_t evolves according to the following equation

$$\frac{d}{dt} Q_t = \begin{cases} I_t - d : \text{if } Q_t > 0 \text{ or } I_t \geq d \\ 0 : \text{otherwise} \end{cases} \tag{57}$$

It should be noted that Q_t is a deterministic function of the sample path of I_t, but it is useful to consider (Q_t, I_t) as a variate Markov process representing the state of the system. Now introduce the integrated process as follows

$$Q_t = q + \int_0^t I_s ds - \int_0^t d * dt = q + A_t - D_t \tag{58}$$

where q is the buffer content at time 0, A_t is the total arrival rate till time t and D_t is the total flow out of the system till time t. A_t is already modelled as a integrated diffusion process and we will now approximate the output process D_t also as a diffusion process. Consider the second integral alone. This integral can be represented as

$$D_t = \int_0^t I(Q_s > 0) d * ds + \int_0^t I(Q_s = 0) I_s ds. \tag{59}$$

The first term represents that the output rate is equal to d when the buffer content is greater than zero or when the instantaneous arrival rate is greater than the instantaneous departure rate. The second term represents the fact that the output rate will be equal to the input rate when the instantaneous arrival rate is less than the departure rate and the buffer content is zero. We must evaluate the mean $E[I_s/I_s < d]$ and variance $\mathrm{var}[I_s/I_s < d]$ of the second term. In order to evaluate the second term, we approximate I_t on $\{Q_t = 0\}$ by an one dimensional diffusion process

$$\hat{\mu} dt + \omega d\hat{W}_t \tag{60}$$

where $\hat{\mu} = E[I_s / I_s < d]$ and $\omega = \sqrt{\mathrm{var}[I_s / I_s < d]}$. The derivation of the conditional mean and variance is given in Appendices A.3 and A.4. So we are approximating the output process by a diffusion process. The ω need not be evaluated because it is not relevant to the approximations of the steady-state probabilities of Q_t and its reasons will be discussed in the next section. Substituting the above values in eqn (58), we get the approximating diffusion equation for the queue length process as

$$X_t = x + A_t - D_t$$

$$= x + \int_0^t d(\mu s + \sigma W_s) - \int_0^t I(X_s > 0) d * ds - \int_0^t I(X_s = 0) d(\hat{\mu} s + \omega W_s)$$

$$X_t = q + \int_0^t I(X_s > 0) d[(\mu - d)s + \sigma\hat{W}s] - \int_0^t I(X_s = 0) d[(\mu - \hat{\mu})s + \omega W_s]$$

where W_s and $\hat{W}_s$ are the standard Wiener processes. Both W_t and $\hat{W}_s$ are Wiener processes and in future can be represented by just W_t. Since the sets $\{X_t > 0\}$ and $\{X_t \le 0\}$ are disjoint, the above equation can be written as the following diffusion equation

$$dX_t = \begin{cases} (\mu - d)dt + \sigma dW_t \ : \ X_t > 0 \\ (\mu - \hat{\mu})dt + \omega dW_t : \ X_t \le 0 \end{cases} \tag{62}$$

The approximating diffusion process derived for the queue length is derived after several steps of approximations. Therefore theoretically there is no guarantee that the process Q_t simulates the original behaviour of the original queue length process when it is small. But for large queue lengths it may well simulate because of the following reason. let the queue length at time t_0 be q_{t_0} and let τ be the last departure time from level 0 of the queue length process before time t_0. During the interval $(\tau, t_0]$, the queue length process is greater than 0 and the departure rate process is d. Thus we have

$$X_t = A_t - A_\tau + d(t-\tau) \text{ for } t \in (\tau, t_0]. \tag{63}$$

This approximation is justified for large $t - \tau$.

4.2. *Stationary solution*

The above equation satisfies the standard Fokker–Planck equation

$$\frac{\partial P(t,x)}{\partial t} = -\frac{\partial}{\partial x} P(t,x)a(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} P(t,x)b(x). \tag{64}$$

At steady state, the above equation satisfies

$$0 = -\frac{\partial}{\partial x} P(t,x)a(x) + \frac{1}{2}\frac{\partial^2}{\partial x^2} P(t,x)b(x). \tag{65}$$

Solving the above equation, we get

$$P(x)\frac{A}{b(x)} e^{2\int_0^x \frac{a(y)}{b(y)}dy} \tag{66}$$

where the constant A is found from

$$A^{-1} = \frac{1}{b(x)}\int_0^\infty e^{2\int_0^q \frac{a(y)}{b(y)}dy} dq \tag{67}$$

So

$$P(x) = P_+(x) \text{ for } X_t > 0 \text{ and } P_-(x) \text{ for } X_t \leq 0 \tag{68}$$

where

$$P_+(x) = \frac{A}{\partial^2} e^{2\frac{\mu-d}{\sigma^2}x} \quad \text{for } X_t > 0, \text{ and} \tag{69}$$

$$P_-(x) = \frac{A}{\omega} e^{2\frac{(\mu-\hat{\mu})}{\omega^2}} \text{ for } X_t \leq 0. \tag{70}$$

The constant A is given by

$$A^{-1} = \frac{1}{\sigma}\int_0^\infty e^{2\frac{\mu-d}{\sigma}x} dx + \frac{1}{\omega}\int_{-\infty}^0 e^{2\frac{\mu-\hat{\mu}}{\omega}} dx = \frac{d-\hat{\mu}}{2(d-\mu)(\mu-\hat{\mu})}. \tag{71}$$

From this it is clear that $P(x)$ is independent of ω for $q > 0$ and the boundary condition at $q = 0$ should be set as $\sigma P(0^+) = \sigma P(0^-)$. Denote Q as the stationary distribution of the random variable Q_t. Now $P[Q > x]$ can be approximated as

$$P[Q > x] \approx \int_x^{\infty} P(u)du \tag{72}$$

$$= L\int_x^{\infty} e^{-2\frac{d-\mu}{\sigma^2}u} du \tag{73}$$

$$= Le^{-2\left(\frac{d-\mu}{\sigma^2}\right)x}. \tag{74}$$

The stationary distribution of X_t is an approximation for Q, the stationary random variable of original queue length process Q_t and the mass at $x = 0$ is approximated as $P(Q = 0) \approx 1 - L$, where $L = \frac{\mu - \hat{\mu}}{d - \hat{\mu}}$.

The above approximation, for the queue length exceeding a particular threshold x, is similar to the one obtained by Simonian[6] and Kobayashi and Ren.[7] Figures 10(a) and (b) give the relationship of the buffer size with cell-loss probability. The first figure is in the range of 700 to 1200 cells and the second covers for the buffer size between 1000 and 1500 cells. The queue-length process is an approximating process and there is no guarantee that the process simulates the original behaviour of the Q_t when it is small. However, this process approximates Q_t for large t.

We make comparisons with that of Anick *et al.*[11] and Simonian[6] for the cell-loss probability. Here, we consider $\log_{10}P(Q > x)$ instead of just $P(Q > x)$. First, we consider the heavy traffic case in which the input is close to the output.

It is very clear that our model gives the worst case performance for all values of the buffer size. So our model acts as the upper bound on the cell-loss probability. This is also validated by testing the spring sequence. We use moment matching techniques and determine the parameters like mean μ and variance σ for our diffusion model and compare our analytical model

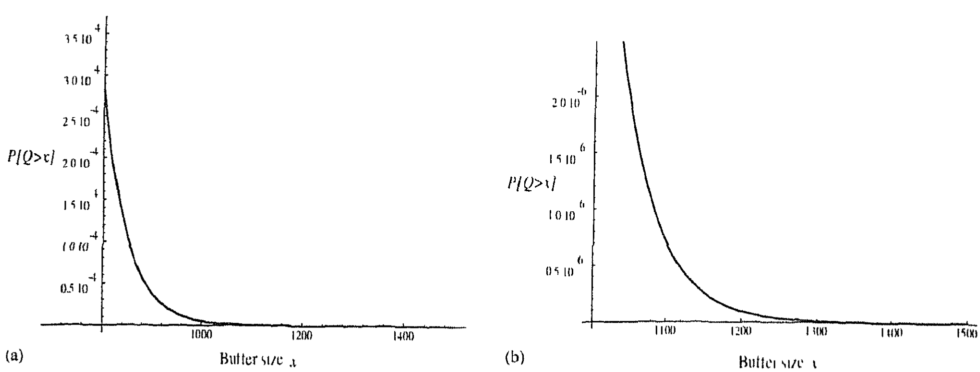

FIG. 10. Cell-loss probability vs buffer size in cells (ranges (a) 700–1200 and (b) 1000–1500).

Tabel I
Comparison of cell-loss probability, at heavy traffic condition, among Anick's Simonian's and our models for different values of buffer sizes

B (cells)	Anick model	Simonian Model	Our model
	$\log_{10}P_L$	$\log_{10}P_L$	$\log_{10}P_L$
0	–0.37	0	–0.07432
2	–0.82	–0.28	–0.178018
4	–1.11	–0.56	–0.356036
6	–1.39	–0.85	–0.534054
8	–1.65	–1.13	–0.712072
10	–1.92	–1.41	–0.89009
20	–3.22	–2.82	–1.78018
30	–4.52	–4.23	–2.67027
40	–5.82	–5.64	–3.56036
50	–7.12	–7.05	–4.45045
60	–8.42	–8.46	–5.34054
70	–9.72	–9.87	–6.23063
80	–11.02	–11.28	–7.12072

Here we assume the output rate to be 16.66 mean(μ) = 14.29, $\sigma = 3.19$ and $\eta = 1.4$.

with the original trace and Simonian model (Tables I to III). The tables show that our model achieves the upper bounds on the cell-loss probability.

Now consider the case when the input is less than that of the output. This is the case of mild traffic. The output rate is equal to 16.66, mean(μ) = 7.14, $\sigma = 2.26$ and $\eta = 1.4$. Our model performs very badly in this case as is the case with usual fluid models, since fluid models are not expected to behave well under mild traffic conditions. The diffusion process for the queue

Tabel II
Comparison of cell-loss probability, between original trace sequence of the spring, Simonian's and our model for different values of buffer sizes

B (cells)	Original Trace	Simonian model	Our model
	$\log_{10}P_L$	$\log_{10}P_L$	$\log_{10}P_L$
0	–0.37	0	–0.01
100	–0.51	–0.21	–0.14
200	–0.69	–0.42	–0.30
300	–0.75	–0.63	–0.44
400	–0.83	–0.84	–0.60
450	–0.88	–0.95	–0.67
500	–0.93	–1.06	–0.74
550	–0.98	–1.16	–0.82
600	–1.05	–1.27	–0.90

The output rate is 250, mean(μ) = 215, $\sigma = 100.83$ and $\eta = 1.4$ are obtained using moment-matching techniques.

Tabel III
Comparison of cell-loss probability, at light traffic condition, among Anick's, Simonian's and our model for different values of buffer sizes

B (cells)	Anick model	Simonian model	Our model
	$\log_{10}P_L$	$\log_{10}P_L$	$\log_{10}P_L$
0	−4.13	−8.72	−1.0034
2	−8.14	−10 99	−2.00578
4	−10.56	−13.25	−4.01157
6	−12.77	−15.52	−6.01735
8	−14.91	−17.78	−8.02313
10	−17.02	−20.05	−10.0289

The output rate is equal to 16.66, mean(μ) = 7.14, σ = 2.26 and η = 1.4.

length process has been obtained after several steps of approximations. Therefore theoretically there is no guarantee that the queue length process will simulate the original process. But Q_t simulates the original process when t is large. Suppose that $Q_{t_0} > 0$ and let τ be the last departure time from level 0 of the process Q_t before time t_0. During this time interval $(\tau, t_0]$, $Q_t > 0$ and the departure rate is d. Thus we have

$$Q_t = A_t - A_\tau + d(t - \tau) \text{ for } t \in (\tau, t_0] \quad (75)$$

So if the output rate is d, then we have the queue length greater than 0 and so the diffusion process may simulate the original process if the interval $(\tau, t_0]$ is large.

Conclusion

In this paper, we have developed diffusion models for the video telephony traffic. We have taken a simple Markov chain interpreted for the slow moving video as in Maglaris *et al.*[3] and mapped it into a truncated diffusion process. Buffer behaviour for such sources is also derived. Our results of the plots of probability density functions and comparison with Anick *et al.*[1] and Simonian[6] show that our model gives a reasonable approximation for Video Telephony. We intend to improve upon this model by incorporating the scene changes thus extending the scope to video traffic in general.

References

1. Anick, D., Mitra, D. and Sondhi, M. M. Stochastic theory of data-handling system with multiple sources, *Bell Syst. Tech J.*, 1982, **8**, 1871–1894

2. Elwalid, A. L. and Mitra, D. Analysis and design of rate based congestion control of high speed networks I:-Stochastic fluid models, Access regulation, *Queueing Systems*, 1991, **9**, 29–64.

3. Maglaris, B., Anastassiou, D., Sen, P., Karlsson, G. and Robbins, J. D. Performance models of statistical multiplexing in packet video communications, *IEEE Trans. Commun.*, 1988, **36**, 834–844.

4. Karandikar, R. L. and Kulkarni, V. G. Second-order fluid flow models: Reflected Brownian motion in a random environment, *Op. Res.*, 1995, **43**, No. 1, 77–88.

5. FELLER, W. An introduction to probability theory and its applications, vol. 2, Wiley, March 1991.

6 SIMONIAN, A. Stationary analysis of a fluid queue with input rate varying as an Ornstein-Uhlenbeck process, *SIAM J Appl. Math.*, 1991, **51**, 828–842.

7 KOBAYASHI, H AND REN, Q A diffusion approximation analysis of an ATM statistical multiplexer with multiple types of traffic, Part I-Equilibrium state solutions, *Proc. IEEE Conf on Commun.*, May 1993, vol. 2, pp. 1047–1053.

8 BLONDIA, C AND CASALS, O. Statistical multiplexing of VBR sources: A matrix-analytic approach, *Per Eval*, 1992, **16**, 5–20

9. HEYMAN, D. P AND LAKSHMAN, T. V. Source models for VBR broadcast-video traffic, *IEEE/ACM Trans Networking*, 1996, **4**, No 1, 40–48.

10. LOU, W., CHIA, L T. AND LEE, B. S. Characterization and source modeling of MPEG-2 VBR video source, *Proc. First Int. Conf on Inform. Commun and Signal Proc*, Singapore, 1997, vol. 2, pp 1652–1656

11. LUCANTONI, D. M New results on the single server queue with the batch Markovian arrival process, *Stoc Mod.*, 1991, **7**, 1–46.

12. COX, D R. AND MILLER, H D *The theory of stochastic processes*, Chapman and Hall, 1977.

13 WONG, E. *Stochastic processes in information and dynamical systems*, McGraw-Hill, 1971

14. GELENBE, E. On approximate computer system models, *J. ACM.*, 1975, **22**, 261–269.

15. SCHWARTZ, M. *Broadband integrated networks*, Prentice Hall, 1996.

16 SOMMERFELD, A. *Partial differnetial equations in physics*, Academic Press, 1949

17 MCKEAN, H P. *Stochastic integrals*, Academic Press, 1969.

18 RESNICK, S *Adventures in stochastic processes*, Birkhauser, 1992.

19 SHWARTZ, A. AND WEISS, A. *Large deviations for performance analysis-Queues, communica tions and computing*, Chapman and Hall, 1996.

Appendix

A. Derivation of diffusion equation from discrete-time Markov chain (DTMC)

Squaring eqn (7) and adding to eqn (8), we get

$$[P_{i,i+1} + P_{i,i-1}]\Delta^2 = \mu^2(i\Delta)\Delta t + \sigma^2(i\Delta)\,(\Delta t)^2 \tag{A1}$$

Using eqns (7) and (A.1), we get

$$P^{\Delta}_{i,i+1} = \frac{\mu(i\Delta)\Delta t}{2\Delta} + \frac{\sigma^2(i\Delta)\Delta t + \mu^2(i\Delta)\Delta t}{2\Delta^2} \tag{A2}$$

and

$$P^{\Delta}_{i,i-1} = -\frac{\mu(i\Delta)\Delta t}{2\Delta} + \frac{\sigma^2(i\Delta)\Delta t + \mu^2(i\Delta)\Delta t}{2\Delta^2}. \tag{A3}$$

Now consider the transition from state i to any arbitrary state j in exactly $n+1$ steps. The equation governing such a transition is given by

$$P_{i,j}^{n+1} = P_{i,i+1}P_{i+1,j}^{n} + P_{i,i}P_{i,j}^{n} + P_{i,i-1}P_{i-1,j}^{n}. \tag{A4}$$

Substituting the values of $P_{i,i+1}$, $P_{i,i}$ and $P_{i,i-1}$ in eqn (A4) we get

$$\left[\frac{\mu(i\Delta)\Delta t}{2\Delta}+\frac{\sigma^2(i\Delta)\Delta t+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i+1,j}^{n}+2\left[\frac{1}{2}-\frac{\sigma^2(i\Delta)\Delta t+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i,j}^{n}+$$

$$\left[-\frac{\mu(i\Delta)\Delta t}{2\Delta}+\frac{\sigma^2(i\Delta)\Delta t+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i-1,j}^{n}. \tag{A5}$$

Taking $P_{i,j}$s on one side we get

$$\frac{P_{i,j}^{n+1}-P_{i,j}^{n}}{\Delta t}=\left[\frac{\mu(i\Delta)}{2\Delta}+\frac{\sigma^2(i\Delta)+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i+1,j}^{n}+2\left[\frac{\sigma^2(i\Delta)+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i,j}^{n}+$$

$$\left[-\frac{\mu(i\Delta)}{2\Delta}+\frac{\sigma^2(i\Delta)\Delta t+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]P_{i-1,j}^{n} \tag{A6}$$

$$\frac{\mu(i\Delta)}{2\Delta}\left[P_{i+1,j}-P_{i-1,j}\right]+\left[\frac{\sigma^2(i\Delta)+\mu^2(i\Delta)\Delta t}{2\Delta^2}\right]\left[P_{i+1,j}^{n}-2P_{i,j}^{n}+P_{i+1,j}^{n}\right] \tag{A7}$$

Equation (A7), when simplified yields

$$\frac{\partial P(t;x,y)}{\partial t}=\mu(x)\frac{\partial P(t;x,y)}{\partial x}+\frac{\sigma^2(x)}{2}\frac{\partial^2 P(t;x,y)}{\partial x^2}.T>0.0<x,y<M. \tag{A8}$$

II. *Derivation of conditional probability density function in the truncated case using the method of images*

Consider a standard Fokker-Planck equation with the boundary condition

$$\left[\frac{1}{2}\sigma^2\frac{\partial f}{\partial x}-\mu f\right]_{x=0}=0. \tag{A9}$$

The appropriate image system for the above problem as in Sommerfeld[15] consists of a point image at $x=-x_0$ and a continuous system of images in the range $x<-x_0$. We get

$$f(x;t)=\frac{1}{\sqrt{2\pi t}\sigma}\left[e^{\frac{-(x-x_0-\mu t)^2}{2\sigma^2 t}}+Ke^{\frac{-(x-x_0-\mu t)^2}{2\sigma^2 t}}+\int_{-\infty}^{x_0}e^{\frac{-(x-a-\mu t)^2}{2\sigma^2 t}}L(a)da\right] \tag{A10}$$

The above equation satisfies the standard Fokker-Planck equation and its initial condition. We must determine the constant K and the function $L(a)$ so that the eqn (A9) is satisfied. Using eqn (A10), the condition eqn (A9) leads to the equations

$$\frac{1}{2}\sigma^2\frac{dL}{da} - \mu L(a) = 0 \text{ and} \tag{A11}$$

$$(x_0 - \mu t)e^{-\frac{x_0\mu}{\sigma^2}} - \left[x_0K + \mu tK + t\sigma^2 L(-x_0)\right]e^{\frac{x_0\mu}{\sigma^2}} = 0, t > 0. \tag{A12}$$

It follows that

$$K = e^{-\frac{2x_0\mu}{\sigma^2}}, L(a) = \frac{2\mu}{\sigma^2}e^{-\frac{2\mu a}{\sigma^2}}. \tag{A13}$$

Substituting the above values in eqn (A10), we get the solution for the Fokker-Planck equation which is bounded at one end.

AIII. Computation of $E[I_s/I_s < d]$

The $P[I_s < a/I_s < d]$ is defined only for $a < d$ as it is zero for $a > d$. We consider a general normal distribution as I_s is defined to be a diffusion process which has the property of being normally distributed (Since I_t can be expressed by forward equation as in Cox and Miller,[12] we have the conditional probability given by normal distribution). Then we consider the truncated normal which is the realistic assumption as the departure and the arrival rates can never be negative in the context of computer communication.

$$\hat{\mu} = E[I_s / I_s < d] = \frac{\int_{-\infty}^{d} xe^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}dx}{\int_{-\infty}^{d} e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}dx} \tag{A14}$$

This expression is got from the fact that $E(x) = E(x)/P(x < d)$. The above equation is a solution for the process which is not restricted, i.e., it is in the range from $(-\infty, +\infty)$.

$$\hat{\mu} = E[I_s / I_s < d] = \frac{\int_{-\infty}^{d} xe^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}dx}{\int_{-\infty}^{d} e^{-\frac{(x-\mu t)^2}{2\sigma^2 t}}dx} \tag{A16}$$

Since μ, σ and d are fixed, we can evaluate $\hat{\mu}$ easily.

AIV. Computation of $\text{var}(I_s/I_s < d)$

Once the mean is found, the variance can be easily calculated from the formulae

$$\text{var}(I_s/I_s < d) = E[(I_s - E(I_s/I_s < d))^2/I_s < d] \tag{A17}$$