# The BRAVO effect and other problems involving ‘biological’ models

D. J. Daley

*Department of Mathematics and Statistics*
*The University of Melbourne*

*[ALSO: The Australian National University, Canberra]*

Conference in Honour of Søren Asmussen

New Frontiers in Applied Probability

Sandbjerg, 1–5 August 2011

SA–DD: Two-sex Galton–Watson branching proc. (*ZW*, 1968)

[DD proved 'obvious' sufficient condition for a.s. extinction using complex variable technique; SA gave martingale proof]

Visit to Australia c.1980 or 1983 ? (Pat Moran's office, view of lake).

etc.

Overlapping visits in Santa Barbara 1988

Oberwolfach meetings . . .

Mittag-Leffler meeting c.2004

AP Editor-in-Chief

Host on briefer visits: Goteborg '91, Aalborg '94, Aarhus '05 ([pipe] organ)

New Frontiers in AP

## 1. A digression (?)

*Epidemics and Rumours in Complex Systems*

Moez Draief and Laurent Massoulié (Cambridge UP, 2010)

Basically about Graph Theory applicable to spreading processes in models for epidemics, rumours (information spread)

Two parts: network unstructured or structured

Counting problems

Math'l techniques giving 'solutions' (martingales, Chernoff bounds [ex Tchebychef inequality])

Connection between microscopic (stochastic) models and macroscopic (deterministic) models
(d.e. methods for latter — Kurtz' theorem)

Graph-theoretic ideas: to what extent are they applicable to (locally finite) infinite stochastic models (on $\mathbb{R}^d$) ??

[Population processes that remain locally finite ??]
(percolation in germ–grain models — Gunter Last . . . )

Yoni Nazarathy and Gideon Weiss (QUESTA 2008)

BRAVO effect:

**B**alancing **R**educes **Asymptotic Variance** of **O**utputs

$$\operatorname{var} N_{\text{dep}}(0,t]$$

M/M/1/K, Buffer of size $K$, Stationary

Arrivals are Poisson at rate $\lambda$,

Service times i.i.d. exponential at rate $\mu$,

With $\rho = \lambda/\mu$ and $t \to \infty$,

$$\operatorname{var} N(0,t] \sim \begin{cases} \rho t & \text{if } \rho < 1, \\ t & \text{if } \rho > 1, \\ \frac{2}{3}\, t & \text{if } \rho \approx 1, \end{cases}$$

$$\text{because output} \approx \begin{cases} \text{input} & \text{if } \rho < 1, \\ \text{max service rate} & \text{if } \rho > 1. \end{cases}$$

[NW08] figures

What happens when $\rho \approx 1$ ?

What happens in many-server system ?

What if reneging or abandonment in place of buffer ?

(joint work with Yoni Nazarathy)

For a stationary orderly point process $N$,

$$\operatorname{var} N(0,t] = \int_0^t \big(2[U(u) - mu] - 1\big)\, m\,\mathrm{d}u,$$

where $U(u) = \mathrm{E}\big[N[0,u] \mid N(\{0\}) > 0\big]$ and $m = \mathrm{E}(N(0,1]$ .

For a renewal process, $U$ is renewal function, and if generic lifetime $X$ has finite second moment, then

$$\operatorname{var} N(0,t] \sim \frac{\mathrm{E}(X^2)}{[\mathrm{E}(X)]^2}\,\frac{t}{\mathrm{E}(X)}$$

If further $E(X^3) < \infty$, then exact linear asymptotics hold i.e.

$$\operatorname{var} N(0,t] = At + B + o(1)$$

for finite constants $A$ and $B$; renewal-theoretic arguments suffice.

Both these properties hold for Markov renewal processes with finite second or third moments.

Queueing O/P in general not Markov renewal, let alone renewal . . .

Output = Arrivals – lost customers

CONSERVATION arguments.

e.g. $k$-server system and $K$ waiting places:

$Q(t)$ = stationary number of customers in system.

$$Q(0) + N_{\mathrm{adm}}(0,t] = N_{\mathrm{dep}}(0,t] + Q(t)$$

$$|N_{\mathrm{adm}}(0,t] - N_{\mathrm{dep}}(0,t]| \le k + K$$

$$\operatorname{var}\left(\frac{N_{\mathrm{dep}}(0,t]}{\sqrt{t}} - \frac{N_{\mathrm{adm}}(0,t]}{\sqrt{t}}\right) \to 0 \qquad (t \to \infty)$$

**Theorem.** *In a stationary G/G/k/K queueing system for which* $\operatorname{var} N_{\mathrm{arr}}(0,1) < \infty$, *the limits as* $t \to \infty$ *of*

$$\frac{\operatorname{var} N_{\mathrm{dep}}(0,t]}{t} \quad \textit{and} \quad \frac{\operatorname{var} N_{\mathrm{adm}}(0,t]}{t}$$

*either both exist finite and are equal, or both are infinite.*

(Sufficient condition for crude asymptotic linearity.)

Turn to *detailed* conservation equations for point processes (sample path realizations — cf. Bremaud (1981) ).

[NW08] is about M/M/1/$K$ (and M/M/$k$/$(K-k)$). $N_{\text{dep}}$ is NOT renewal for $K \geq 2$ but the refined limit behaviour holds for M/M/$k$/$K$ because $Q(t)$ is finite state space continuous time Markov chain and asymptotics for geometrically ergodic chains apply.

A 'quick' route to expressions for the moment behaviour of $N_{\text{dep}}$ comes from point process expressions, using $N_{\text{arr}}$ and $N_{\text{serv}}$ to describe counting functions of arrival point processes and *potential* service departure epochs:

Use $I_j(t)$ to denote an indicator function for $\{Q(t) = j\}$: then in M/M/1/$K$,

$$N_{\text{lost}}(0,t] := \int_{(0,t]} I_{1+K}(u-)\, N_{\text{arr}}(\mathrm{d}u),$$

hence

$$N_{\text{adm}}(0,t] = \int_{(0,t]} \big(1 - I_{1+K}(u-)\big)\, N_{\text{arr}}(\mathrm{d}u).$$

Similarly

$$N_{\text{dep}}(0,t] := \int_{(0,t]} \big(1 - I_0(u-)\big)\, N_{\text{serv}}(\mathrm{d}u).$$

Taking expectations appropriately, e.g.

$$\begin{aligned}
&\mathrm{E}\big[\big(N_{\text{adm}}(0,t]\big)^2\big] \\
&= \mathrm{E}\Big[\int_{(0,t]\times(0,t]} [1 - I_{1+K}(u-)][1 - I_{1+K}(v-)] N_{\text{arr}}(\mathrm{d}u)\, N_{\text{arr}}(\mathrm{d}v)\Big]
\end{aligned}$$

This leads ultimately to

$$\begin{aligned}&\operatorname{var} N_{\mathrm{adm}}(0,t] - \mathrm{E}(N_{\mathrm{adm}}(0,t])\\&\qquad = 2\lambda\mu\pi_0 \int_0^t \big[(t-u)(p_{0,1+K}(u) - \pi_{1+K})\big]\,\mathrm{d}u.\end{aligned}$$

The coarse asymptotics follow by extracting a factor $t$ and then standard convergence property of the integral.

To extract the fine asymptotics, write integral as

$$t\int_0^\infty [p_{0,1+K}(u)-\pi_{1+K}]\,\mathrm{d}u - \int_0^\infty u[p_{0,1+K}(u)-\pi_{1+K}]\,\mathrm{d}u + o(1)$$

where the $o(1)$ term takes account of the discrepancy between the finite and infinite integration, and the other terms have finite limits because of geometric ergodicity and monotonicity of the transition probability functions.

*This technique for studying O/P works for M/M/k/K*

What are implications for graph $Q(t)$ $v$. $t$ ?

(ditto) BRAVO effect ?

(ditto) both of the above for M/M/$k$/rneg

A PROBLEM

The departure process $N_{\text{dep}}$ of these M/M/$k/K$ systems is certainly not renewal, though it is irreducible Markov renewal. As a point process, there is embedded in it a sequence of regenerative epochs:

*What can be said about limit properties of a point process containing an embedded regenerative structure*?

(think of variance behaviour (!) )

For a stationary renewal process, the fine detailed asymptotics hold as soon as the lifetime distribution has a third moment. Do these carry over to a stationary point process that contains an embedded regenerative structure ?

Refer to integral for variance:

$$\operatorname{var} N(0,t] = mt + 2\int_0^t [U(u) - mu]\,\mathrm{d}u$$

Depends of rate of convergence of $U(u) - mu$ to its limit (if it exists) (for renewal process, limit $= \mathrm{E}(X^2)/2[(E(X))^2]$; renewal theorem does not yield full detail of convergence rate, though finite third moment does yield finiteness on

$$U(u) - mu - \tfrac{1}{2}(\text{approx'n to 2nd moment}).$$

OUTPUT OF M/M/$k$/$K$

Recall: $\{\pi_i\}$ is stationary queue-size distribution,

$$\pi_i = \Pr\{Q(t) = i\} \qquad \text{(all } t\text{)}$$

$$\lambda \pi_{i-1} = \min(i, k)\, \mu \qquad (i = 1, \ldots, k, \ldots, k + K),$$

$\sum_{i=0}^{k+K} \pi_i = 1$. For BRAVO effect, want arrival and service rates around 'balance', i.e. $\mu = k\lambda$. Recurrence relations give

$$\pi_i = \begin{cases} \dfrac{(k\lambda/\mu)^i}{i!}\pi_0 \;=\; \dfrac{(k\rho)^i}{i!}\pi_0 & \text{for } i \le k, \\ (\lambda/\mu)^{i-k}\pi_k \;=\; \rho^{k-i}\pi_k & \text{for } i \ge k, \end{cases}$$

First investigate case $\rho = 1$:

Cases $i \le k$ give $\displaystyle\sum_{i=0}^{k} \pi_i \approx \tfrac{1}{2}\mathrm{e}^k \pi_0 \approx \tfrac{1}{2}\pi_k\sqrt{2\pi k}$ for $k$ not small.

Cases $i > k$ give $\displaystyle\sum_{i=k+1}^{k+K} \pi_i = K\pi_k$.

Special case: $\dfrac{\pi_k}{\pi_0} = \dfrac{k^k}{k!} \approx \dfrac{k^k}{\sqrt{2\pi k}\, k^k\, \mathrm{e}^{-k}} = \dfrac{\mathrm{e}^k}{\sqrt{2\pi k}}$ .

Hence, $\quad \pi_k\Big(K + \sqrt{\pi k/2}\Big) \approx 1.$

Want to evaluate crude linear asymptote (this exists, and fine linear asymptotic relation also, because the stationary MC $\{Q(t)\}$ has finite state space and is irreducible, hence it is geometrically ergodic).

Introduce the family of indicator random variables $J_{Q(t-)}$ which, conditional on $Q(t-)$, are mutually independent for distinct time variables $t$ and independent of $N_{\mathrm{serv}}(\mathrm{d}u)$ in $u \geq t$, for which

$$J_{Q(t-)}\big|\{Q(t-) = i\} = \begin{cases} 1, & \text{with probability } \min(i,k)/k, \\ 0, & \text{otherwise.} \end{cases}$$

Then $N_{\mathrm{dep}}(\mathrm{d}t) = J_{Q(t-)}\, N_{\mathrm{serv}}(\mathrm{d}t)$, equivalently

$$N_{\mathrm{dep}}(0,t] = \int_{(0,t]} N_{\mathrm{dep}}(\mathrm{d}u) = \int_{(0,t]} J_{Q(u-)}\, N_{\mathrm{serv}}(\mathrm{d}u).$$

This leads ultimately to

$$\begin{aligned}
&\operatorname{var} N_{\mathrm{dep}}(0,t] - \mathrm{E}\big(N_{\mathrm{dep}}(0,t]\big) \\
&= \frac{2\mu^2}{k^2} \sum_{i=1}^{k+K} \sum_{j=1}^{k+K} \min(i,k)\, \min(j,k) \int_0^t (t-u)\pi_i[p_{i-1,j}(u) - \pi_j]\,\mathrm{d}u, \\
&= \frac{2\mu\lambda}{k} \sum_{j=1}^{k+K} \min(j,k) \int_0^t (t-u)\pi_{k+K}[\pi_j - p_{k+K,j}(u)]\,\mathrm{d}u,
\end{aligned}$$

so

$$\begin{aligned}
&\lim_{t\to\infty} \big[\operatorname{var} N_{\mathrm{dep}}(0,t] - \mathrm{E}\big(N_{\mathrm{dep}}(0,t]\big)\big]/\lambda t \\
&= \frac{2\mu}{k} \sum_{j=1}^{k+K} \min(j,k) \int_0^\infty \pi_{k+K}[\pi_j - p_{k+K,j}(u)]\,\mathrm{d}u,
\end{aligned}$$

and exploiting reversibility, this equals

$$\begin{aligned}
&\frac{2\mu}{k} \sum_{j=1}^{k+K} \min(j,k)\pi_j \int_0^\infty [\pi_{k+K} - p_{j,k+K}(u)]\,\mathrm{d}u \\
&= 2\lambda \sum_{j=1}^{k+K} \pi_{j-1} \int_0^\infty [\pi_{k+K} - p_{j,k+K}(u)]\,\mathrm{d}u.
\end{aligned}$$

[Now convert this to sums of moments of first-passage times etc.]

What changes for $\rho \neq 1$ ?

Use $\rho = 1 - \dfrac{\beta}{K}$, and $K = \alpha\sqrt{k}$

(so both $k,\ K \to \infty$ but 'controlled' relative rate).

Use

$$\lim_{t\to\infty} \frac{\operatorname{var} N_{\mathrm{dep}}(0,t]}{\mathrm{E}\big(N(0_{\mathrm{dep}}(0,t]\big)} = 1 - 2\sum_{i=0}^{k+K} \pi_i v_i (1 - v_i)$$

where (birth-and-death process) $v_i = \pi_{k+K} P_i / \pi_i$ and $P_i = \sum_{j=0}^{i} \pi_i$.

$\operatorname{var} N_{\mathrm{dep}}(0,t]$ falls short of asymptotic rate for Poisson process

$$x(1-x) \leq \frac{1}{4} \qquad \text{(all real } x\text{)}$$

[BUT: expression is for finite state-space birth–death proc.]

Return to [NW08] figure:

It is *variance* function that is asymptotically discontinuous . . . there is change in mechanism producing the O/P process at $\rho = 1$: why should second-order (variability) effect remain continuous like first-order (mean) effect?

Why should volatility near change point be 'continuous' ?

Draief and Massoulié emphasize the criticality theorem for branching processes as simplest change-point phenomenon with small change in reproduction rate produces catastrophic change in ultimate population.

Does BRAVO effect change output from more complex queueing systems as system nears cricality ('heavy traffic') ?