Convergence of Independent Particle Systems

by

John R. Hoffman*

and

Jeffrey S. Rosenthal**

(March 18, 1993; revised August 10, 1994.)

Summary. We consider a system of particles moving independently on a countable state space, according to a general (non-space-homogeneous) Markov process. Under mild conditions, the number of particles at each site will converge to a product of independent Poisson distributions; this corresponds to settling to an ideal gas. We derive bounds on the rate of this convergence. In particular, we prove that the variation distance to stationarity decreases proportionally to the sum of squares of the probabilities of each particle to be at a given site. We then apply these bounds to some examples. Our methods include a simple use of the Chen-Stein lemma about Poisson convergence. Our results require certain strong hypotheses, which further work might be able to eliminate.

Acknowledgements. This line of research was suggested by Fred Hoppe of McMaster University. We are very grateful to him and to Ian Hambleton for their warm hospitality during a recent visit by one of us (JSR). This research was completed at the University of Minnesota, and we wish to thank Larry Gray for invaluable guidance, and Michael Ekhaus and Margaret Short for helpful conversations. Finally, we thank Maury Bramson and Robin Pemantle for a useful discussion of these issues, and thank the referees and editors for very helpful comments.

---

* Secure Computing Corporation, 2675 Long Lake Road, Roseville, MN 55113. Phone: (612) 628-2808. Internet: `hoffman@sctc.com`

** Department of Statistics, University of Toronto, Toronto, Ontario, Canada M5S 1A1. Phone: (416) 978-4594. Internet: `jeff@utstat.toronto.edu`

1. Introduction.

A standard question in Markov process theory is the existence of, and convergence to, a stationary probability distribution. The question of *rate of convergence* concerns how quickly this convergence occurs. Such questions are now standard in the literature (see, e.g. [Di], [DS], [R]).

Many Markov processes do not have normalized stationary distributions, though they may still have a non-negative (but perhaps non-normalizable) invariant measure $m(x)$, $x \in \mathcal{X}$. (We consider only processes on discrete spaces $\mathcal{X}$.) For such processes, we instead consider a *system* of particles, each moving independently according to the Markov process $P^t(x, y)$, for $t \in T$ (in discrete or continuous time). This system process is defined precisely in Section 2, and has been previously studied in [Doo], [Li], and [LP]. Given an invariant measure $m(x)$, the system process will have a stationary probability distribution given by a product over $\mathcal{X}$ of independent Poisson distributions with means $m(x)$. It is thus reasonable to study the question of convergence of the system process to this new stationary distribution.

For such processes, we obtain the following result. The convergence to stationarity, measured by total variation distance on any fixed finite subset K of $\mathcal{X}$, is governed by the quantity

$$s_K(t) \;=\; \sup_{\substack{x \in \mathcal{X} \\ y \in K}} P^t(x, y)\,,$$

the largest single probability of being at a point in K after time t. Specifically, for reasonable initial distributions, the variation distance goes down on the order of between $s_K(t)$ and $s_K(t)^2$. More precise bounds, depending on the subset K and on the initial distributions, are actually presented (Section 3), and it is shown that the distance to stationarity is actually proportional to the sum of squares of certain transition probabilities. Applications to specific examples, including a discussion of the question of exponential convergence rate, are presented in Section 5.

The quantity $s_K(t)$ defined above has been studied in various contexts, including substantial work for certain random walks on groups; see for example [VSC]. Thus, our results can be combined with previous work to get precise information about convergence rates of independent particle systems in such situations.

Our proofs make use of the method of Chen-Stein ([Ch], [Ste]), and in particular the "process version" of Arratia, Goldstein, and Gordon [AGG]. This is discussed in Section 4, along with proofs of our main results.

The convergence to a product of Poissons corresponds to the notion in physics of convergence to an ideal gas; see e.g. [Dy], page 174.

Our work is closely related to a result of Dobrushin [Dob] and Stone [Sto]. They obtain complicated necessary and sufficient conditions for convergence of similar system processes to Poisson point processes. However, their theorem applies only to the space-homogeneous case, where the different particles move according to *translates* of a single Markov process (so in place of $m(x)$ they simply have scaled Lebesgue or counting measure). Furthermore, they give no information about rates of convergence, our main interest here.

In addition, in Section 3 of [De], results are presented regarding the asymptotic decay rate of processes similar to ours, for the case where the underlying Markov chain is a random walk.

Since originally completing this manuscript, we have learned that a similar approach is suggested in Example 10.2.14 of [BHJ]. However, these authors concentrate on the special case of deterministic starting distributions. Furthermore, the bounds they obtain have the awkward property of having transition probabilities in both numerator and denominator, a complication avoided here.

As a running example, consider simple symmetric random walk on the integers $\mathbf{Z}$. This (discrete-time) Markov process is defined by

$$P^{\mathbf{1}}(x, x+1) \;=\; P^{\mathbf{1}}(x, x-1) \;=\; \frac{1}{2}, \qquad x \in \mathbf{Z}$$

with $P^{\mathbf{1}}(x, y) = 0$ otherwise. This Markov process is easily seen to have no invariant probability measure. On the other hand, it is easily seen that counting measure on $\mathbf{Z}$ is an invariant measure for this Markov process, although it is not normalizable. Thus, we will see that the system process induced from this Markov process has a stationary distribution given by a product of independent copies of $Poisson(1)$. Furthermore, here $s_K(t) = O(1/\sqrt{t})$, so that convergence of the system process to stationarity happens on the order of between $1/\sqrt{t}$ and $1/t$. (In fact, we shall prove that it happens on the order of $1/\sqrt{t}$; see Proposition 7.)

2. Preliminaries.

We define the system processes that we wish to study using the following definition, similar to Doob [Doo]. It defines a new Markov process, built out of multiple independent copies of $P^t(x,y)$.

Definition. Let $\{P^t(x,y)\}_{t\in T}$ be the semi-group for a Markov process on a countable state space $\mathcal{X}$. Let ν be a probability measure on ${\mathbf{Z}_+}^{\mathcal{X}}$ (the space of all functions from $\mathcal{X}$ to the set $\mathbf{Z}_+$ of non-negative extended integers). The *(independent) system process* based on $P^t(x,y)$, with initial distribution ν, is the process $\{N_x(t)\}_{t\in T}$ defined on ${\mathbf{Z}_+}^{\mathcal{X}}$ by choosing $\{N_x(0)\}_{x\in\mathcal{X}}$ according to ν, putting $N_x(0)$ particles at each site $x\in\mathcal{X}$, letting them proceed independently according to $P^t(x,y)$, and letting $N_x(t)$ record the number of particles at site x at time t.

By an *invariant measure* for the Markov process $P^t(x,y)$, we shall mean a non-negative (possibly non-normalizable) measure $m(x)$ on the discrete space $\mathcal{X}$, such that

$$\sum_x m(x)\, P^t(x,y) \;=\; m(y)\,, \qquad \text{for all } y\in\mathcal{X}, \quad t\in T\,.$$

The system process defined above allows us to make the connection between a non-normalizable invariant measure as above, and the issue of convergence to a stationary distribution. The connection is given by the following.

Lemma [Doo]. *Let $\{m(x)\}_{x\in\mathcal{X}}$ be an invariant measure for $P^t(x,y)$. Then the measure*

$$\prod_{x\in\mathcal{X}} Poisson\,(m(x))\;,$$

given by a product of independent Poisson distributions with means $m(x)$, is a stationary probability distribution for the system process $\{N_x(t)\}_{t\in T}$ associated with $P^t(x,y)$.

The proof of this lemma is a straightforward. Let $\{N_x(0)\}_{x\in\mathcal{X}}$ be chosen from the above distribution. The invariance of the measure $m(x)$ ensures that the means of $N_x(t)$

will not depend on t. Elementary properties of the Poisson distribution then ensure that the distributions of the $N_x(t)$ will in fact remain as independent Poissons.

The question of whether the system process has *other* stationary distributions, in addition to mixtures of the products of Poissons, is discussed in [Li] and [LP].

The lemma suggests that we study system processes, and consider the question of their convergence (and rate of convergence) to the stationary distribution given above. This is done in the next section.

3. Results.

We consider a Markov process $\{P^t(x,y)\}_{t\in T}$, in discrete or continuous time, on a countable state space $\mathcal{X}$, with a non-negative invariant measure $m(x)$ as above. We shall study the system chain $\{N_x(t)\}_{t\in T}$ defined in Section 2. To study convergence, we shall use the metric given by total variation distance on a *finite subset* of $\mathcal{X}$. (There will usually *not* be convergence in total variation distance on the entire state space.)

Write ν for the initial distribution of $\{N_x(0)\}_{x\in\mathcal{X}}$ on $\mathbf{Z}_+{}^{\mathcal{X}}$. We shall assume that $\nu = \prod\limits_{x\in\mathcal{X}} \nu_x$ is given by a product measure, with ν_x a probability measure on $\mathbf{Z}_+$ having the "correct" finite mean $m(x)$, and having finite second moment $m_2(x)$. This ensures that we will have

$$E\,(N_x(t)) \;=\; m(x), \qquad \text{for all } t\in T\,.$$

We shall prove the following result about convergence to a product of Poissons.

Theorem 1. *Consider the system process $\{N_x(t)\}$ defined above, with $K\subseteq\mathcal{X}$ a finite subset. Let ν, $m(x)$ and $m_2(x)$ be as above, and set $F_K=\min\left(1,\ \max\limits_{x\in K}\left(m(x)^{-1/2}\right)\right)$. Then if the process begins in initial distribution ν, then*

$$\left\| \mathcal{L}\left(\{N_x(t)\}_{x\in K}\right) - \prod_{x\in K} Poisson\,(m(x)) \right\|_{\mathrm{var}} \le 4\,F_K \sum_{x\in\mathcal{X}} \left(m_2(x)+m(x)^2-m(x)\right)\,\left(P^t(x,K)\right)^2\;,$$

where $\|P-Q\|_{\mathrm{var}} = 2\,\sup\limits_{A}|P(A)-Q(A)| = \sum\limits_{x\in\mathcal{X}}|P(x)-Q(x)|$ is the usual total variation distance for probability measures.

Remark. We are unable to obtain a natural lower bound in this case, because of the difficulty of controlling the unknown starting distribution ν. This is improved in Theorem 4 below.

The upper bound given in Theorem 1 may be difficult to interpret. The following simple corollary is perhaps more intuitive.

Corollary 2. *Assume $d \le m(x) \le D$ and that $m_2(x) \le D_2$, for all $x \in \mathcal{X}$. Then*

$$\left\| \mathcal{L}\left(\{N_x(t)\}_{x \in K}\right) - \prod_{x \in K} Poisson\,(m(x)) \right\|_{\mathbf{var}} \le \left(4\, F_K\, |K|^2\, D^2 (D + D_2)/d\right)\, s_K(t)\,,$$

where $|K|$ is the number of elements in K, and where

$$s_K(t) = \sup_{\substack{x \in \mathcal{X} \\ y \in K}} P^t(x, y)\,.$$

Proof. We bound $m(x)^2$ by D^2 and bound $m_2(x)$ by D_2, and note that

$$\sum_{x \in \mathcal{X}} \left(P^t(x, K)\right)^2 \le \left(\sup_{x \in \mathcal{X}} P^t(x, K)\right) \sum_{x \in \mathcal{X}} P^t(x, K)\,.$$

Since $\sum_{x \in \mathcal{X}} m(x) P^t(x, K) = m(K)$, we must have $\sum_{x \in \mathcal{X}} P^t(x, K) \le m(K)/d$. The result now follows from the inequalities $P^t(x, K) \le |K| s_K(t)$ and $m(K) \le |K| D$. ∎

The corollary also asserts that the variation distance will be small if $s_K(t)\,|K|^2$ is small. This shows that for a given t, the process will be approximately independent Poisson on sets K whose size is small compared to $O(1/\sqrt{s_K(t)})$. (Note that Theorem 1 is a more refined statement; the bound there depends on the layout of K, not merely on its size.)

As a further corollary, we can immediately obtain information about weak convergence of our system process in a certain topology. This corollary also follows from Theorem 1.3 of [LP]. The proof is straightforward and is omitted.

Corollary 3. *Assume, in addition to the assumptions of the previous corollary, that*

$$\lim_{t\to\infty} \sup_{x\in\mathcal{X}} P^t(x,y) \;=\; 0, \qquad \textit{for all } y \in \mathcal{X}\,.$$

Then the system process converges weakly to a product of independent Poissons with means $m(x)$, in the usual product topology on ${\mathbf{Z}_+}^{\mathcal{X}}$.

The upper bound given in Theorem 1 works uniformly for *any* initial distributions ν_x with given first and second moments. Under certain additional restrictions on the ν_x, stronger statements can be made. For our next result, we assume the following.

(A1) The distributions ν_x are each given by sums of independent Bernoulli random variables. Specifically, $N_x(0) = \sum_{\gamma\in\Gamma_x} H_\gamma$, where $H_\gamma \sim \textit{Bernoulli}(h_\gamma)$ are independent, where Γ_x is an appropriate index set, and where $\sum_{\gamma\in\Gamma_x} h_\gamma = m(x)$.

Note that these initial distributions include deterministic ones, where $m(x)$ is an integer and $\nu_x(m(x)) = 1$, because it is permissible to have $h_\gamma = 1$.

Assumption (A1) suggests the following interpretation. Let each $\gamma \in \Gamma_x$ represent a distinct particle, which is created at x with independent probability h_γ, or is otherwise not created at all. Then, for $\gamma \in \Gamma_x$, the probability that particle γ was created, and is at site $y \in \mathcal{X}$ at time t, is given by $h_\gamma\, P^t(x,y)$.

Under these conditions, we can prove sharp upper and lower bounds in terms of the quantities

$$J_A(t) \;=\; \sum_{x\in\mathcal{X}}\sum_{\gamma\in\Gamma_x} \big(h_\gamma\, P^t(x,A)\big)^2\,,$$

defined for any subset $A \subseteq \mathcal{X}$, and with $J_y(t) = J_{\{y\}}(t)$. We shall prove the following.

Theorem 4. *Consider the system process $\{N_x(t)\}$ as above. Let K be a finite subset of $\mathcal{X}$, and let $m(x)$, h_γ, F_K, and $J_A(t)$ be as above. Assume that (A1) holds. If the process begins in the initial distribution ν, then*

$$\begin{aligned} 2\max_{y\in K}\left(e^{-m(y)}\left(1 \;-\; e^{-J_y(t)/2}\right)\right) \;&\le\; \Big\| \mathcal{L}\left(\{N_x(t)\}_{x\in K}\right) \;-\; \prod_{x\in K} \textit{Poisson}\,(m(x)) \Big\|_{\mathbf{var}} \\ &\le\; 4\,F_K\,J_K(t)\,, \end{aligned}$$

Note that for small values of $J_y(t)$, the quantity $1 - e^{-J_y(t)/2}$ is approximately $J_y(t)/2$. The theorem thus says essentially that, up to constants independent of t, the variation distance to independent Poissons is bounded between the maximum of $J_y(t)$, and $J_K(t)$. (Note that if $|K| = 1$, then $J_K(t) = J_y(t)$.) Note also that the quantity $\max_{y\in K} e^{-m(y)}$ and the quantity F_K are both bounded above by 1, and both go to 0 as $\min_{y\in K} m(y)$ gets large.

Remark. If the underlying Markov chain is *symmetric*, i.e. $P^t(x,y) = P^t(y,x)$ for all $x, y \in \mathcal{X}$, then there are general bounds on $J_y(t)$ and $J_K(t)$ which may help in applying Theorem 4. Indeed, recalling that $\sum_{\gamma\in\Gamma_x} h_\gamma = m(x)$, we have

$$\begin{aligned} J_y(t) &= \sum_{x\in\mathcal{X}} \sum_{\gamma\in\Gamma_x} \left(h_\gamma P^t(x,y)\right)^2 \ \geq\ (\inf_\gamma h_\gamma)(\inf_x m(x)) \sum_{x\in\mathcal{X}} P^t(y,x)P^t(x,y) \\ &= (\inf_\gamma h_\gamma)(\inf_x m(x)) P^{2t}(y,y)\,. \end{aligned}$$

Also

$$\begin{aligned} J_K(t) &= \sum_{x\in\mathcal{X}} \sum_{\gamma\in\Gamma_x} (h_\gamma)^2 \sum_{k,k'\in K} P^t(x,k)P^t(k',x) \ \leq\ (\sup_\gamma h_\gamma)(\sup_x m(x)) \sum_{k,k'\in K} P^{2t}(k,k') \\ &\leq\ |K|^2 (\sup_\gamma h_\gamma)(\sup_x m(x)) \max_{k,k'\in K} P^{2t}(k,k')\,. \end{aligned}$$

We shall make use of these bounds in proving Proposition 7 below.

As in the case of Theorem 1, we can easily deduce bounds from Theorem 4 which are more intuitive.

Corollary 5. *Assume, in addition to (A1), that $m(x) \leq D$ for all $x \in \mathcal{X}$. Then*

$$\begin{aligned} 3\,(\inf h_\gamma)^2\, e^{-D}\, s_K(t)^2/4 &\leq \left\| \mathcal{L}\{N_x(t)\}_{x\in K} \ -\ \prod_{x\in K} Poisson\,(m(x)) \right\|_{\mathbf{var}} \\ &\leq\ 4\, F_K\, |K|^2\, D\, (\sup h_\gamma)\, s_K(t)\,, \end{aligned}$$

where again $s_K(t) = \sup_{\substack{x\in\mathcal{X}\\ y\in K}} P^t(x,y)$.

Proof. For the lower bound, we note that $e^{-m(y)} \geq e^{-D}$. Also $\max_{y\in K} J_y(t) \geq ((\inf h_\gamma) s_K(t))^2$ by inspection, and $1 - e^{-r} \geq 3r/4$ for $r \leq \frac{1}{2}$, so that $1 - e^{-J_y(t)/2} \geq 1 - e^{-((\inf h_\gamma)s_K(t))^2/2} \geq 3\left((\inf h_\gamma)s_K(t)\right)^2/8$.

For the upper bound, we note that

$$J_K(t) \leq |K|\,(\sup h_\gamma)\, s_K(t)\, m(K) \leq |K|^2\,(\sup h_\gamma)\, D\, s_K(t)\,.$$

The result follows. ∎

This corollary shows that the rate of convergence in this case is essentially governed by the convergence to zero of the quantity $s_K(t)$, which represents the largest single transition probability of the underlying Markov process. In particular, the corollary shows that the convergence of the system process will be exponential if and only if the quantity $s_K(t)$ decreases exponentially as a function of t. This is explored further in Section 5.

In our running example of simple symmetric random walk, it is well-known that $s_K(t) = O(1/\sqrt{t})$, and does not depend on K. Thus, the convergence to stationarity of the associated system process goes down on the order of between $1/\sqrt{t}$ and $1/t$. (In fact, we shall show that it goes down as $O(1/\sqrt{t})$; see Proposition 7.) Also, the system will be approximately independent Poisson on sets of size $o(t^{1/4})$.

Lastly, corresponding to Theorem 4 in the case when $|K| = 1$, one can obtain a quantitative version of the "Law of Rare Events", concerning the approximation of the distribution of sums of binomial random variables (with different means) by a Poisson distribution. (An asymptotic version of this law can be found in [Fe], p. 282.) However, better bounds are already known (see for example [BHJ]), so we omit a precise statement.

Theorems 1 and 2 are proved in the following section. The question of how quickly the quantity $s_K(t)$ decays in specific examples is explored in Section 5.

4. Proof of Theorems 1 and 4.

We shall use the following "process version" of the Chen-Stein method ([Ch], [Ste]), due to Arratia, Goldstein, and Gordon ([AGG], Theorem 4).

Lemma 6. *Let $I = I_1 \cup \ldots \cup I_d$ be an index set, and let $\{X_\alpha\}_{\alpha \in I}$ be a collection of (possibly dependent) indicator variables. For each $\alpha \in I$, choose a neighborhood $B_\alpha \subseteq I$ with $\alpha \in B_\alpha$, and assume X_α is independent of $\{X_\beta\}_{\beta \notin B_\alpha}$. Set $W_j = \sum_{\alpha \in I_j} X_\alpha$, and set $\lambda_j = E(W_j)$. Then*

$$\|\mathcal{L}(W_1, \ldots, W_d) \; - \; \prod_j Poisson(\lambda_j)\|_{\mathrm{var}} \;\leq\; 4 \min\left(1, (\min_j \lambda_j)^{-1/2}\right)(b_1 + b_2),$$

where

$$b_1 = \sum_{\alpha \in I} \sum_{\beta \in B_\alpha} E(X_\alpha)\, E(X_\beta); \qquad b_2 = \sum_{\alpha \in I} \sum_{\substack{\beta \in B_\alpha \\ \beta \neq \alpha}} E(X_\alpha X_\beta).$$

Remark. The original result in [AGG] contained an extra factor of 1.4, but this has been eliminated by Arratia and Tavaré [AT].

We now proceed to the proof of the theorems. Because it is more straightforward, we begin with the proof of Theorem 4.

Proof of Theorem 4. For the lower bound we compute directly. Indeed, for any fixed $y \in K$, we have (writing q_γ for $h_\gamma P^t(x, y)$ for $\gamma \in \Gamma_x$, and writing $\sum_\gamma$ for $\sum_{x \in \mathcal{X}} \sum_{\gamma \in \Gamma_x}$)

$$\log P(N_y(t) = 0) \;=\; \sum_\gamma \log(1 - q_\gamma) \;\;\leq\;\; \sum_\gamma \left[-(q_\gamma) - (q_\gamma)^2/2\right] \;=\; -m(y) - \sum_\gamma (q_\gamma)^2/2,$$

where we have used $\sum_\gamma q_\gamma = m(y)$ and an easy bound on $\log(1-r)$. Hence,

$$(Poisson(m(y);\, 0) \;-\; P(N_y(t) = 0)) \;\geq\; e^{-m(y)}\left(1 \;-\; e^{-\sum (q_\gamma)^2/2}\right).$$

The lower bound follows immediately.

For the upper bound, we use Lemma 6. Write Γ for the disjoint union of the Γ_x, and for $k \in K$, set $I_k = \{(\gamma, k) \,|\, \gamma \in \Gamma\}$, so that $I = \{(\gamma, k) \,|\, \gamma \in \Gamma, k \in K\}$. We choose the neighborhoods $B_{(\gamma,k)} = \{(\gamma, j) \,|\, j \in K\}$. Then the conditions of the lemma are satisfied, for indicator variables $H_{(\gamma,k)}$ defined by $H_{(\gamma,k)} = 1$ if and only if particle γ is at position k at time t. Furthermore, $b_2 = 0$ since the same particle γ can't be at two different sites j and k at the same time. Finally, we compute that

$$b_1 \;=\; \sum_{x \in \mathcal{X}} \sum_{\gamma \in \Gamma_x} \sum_{k \in K} h_\gamma P^t(x, k) \sum_{j \in K} h_\gamma P^t(x, j) \;=\; \sum_{x \in \mathcal{X}} \sum_{\gamma \in \Gamma_x} \left(h_\gamma P^t(x, K)\right)^2.$$

Theorem 4 now follows. ∎

Remark. The upper bound above can also be proved directly, without making use of the Chen-Stein result. However, the proof is more cumbersome and also leads to slightly weaker bounds.

Proof of Theorem 1. We again use Lemma 6. We first realize ν_x as a sum of (dependent) indicator variables. We define indicator random variables $H_{(x,j,k)}$ for $x \in \mathcal{X}$, $j = 1, 2, 3, \ldots$, and $k = 1, 2, \ldots, j$, where $P(H_{(x,j,k)} = 1) = \nu_x(j)$, where $H_{(x,j,k_1)} = H_{(x,j,k_2)}$, and where $(H_{(x,j_1,k_1)})(H_{(x,j_2,k_2)}) = 0$ for $j_1 \neq j_2$. This means that the variable $\sum_{j,k} H_{(x,j,k)}$ is distributed according to ν_x.

We then define indicator variables $H_{(x,j,k,y)}$ by starting independent particles at x corresponding to each $H_{(x,j,k)}$, and setting

$$H_{(x,j,k,y)} = 1 \qquad \text{if and only if} \qquad (H_{(x,j,k)} = 1 \text{ and the particle goes from } x \text{ to } y \text{ in time } t).$$

Thus, $P(H_{(x,j,k,y)} = 1) = \nu_x(j)\, P^t(x, y)$.

We define a neighborhood for (x, j, k, y) given by

$$B_{(x,j,k,y)} = \{(x, j', k', y') \mid \text{any } j', k', y'\}.$$

We claim that

$$b_1 = \sum_x m(x)^2 \left(P^t(x, K)\right)^2 ,$$

and that

$$b_2 = \sum_x (m_2(x) - m(x)) \left(P^t(x, K)\right)^2 .$$

Indeed, since our neighborhoods form an equivalence relation, we can evaluate b_1 by summing over all neighborhoods (indexed by x) the *square* of the sum of the p_α. Thus,

$$\begin{aligned}
b_1 &= \sum_{\substack{\text{distinct}\\ \text{nhbds}}} \left(\sum_{\alpha \in \text{nhbd}} E(X_\alpha) \right)^2 \\
&= \sum_{x \in \mathcal{X}} \left(\sum_{j=1}^{\infty} \sum_{k=1}^{j} \sum_{y \in K} E(H_{(x,j,k,y)}) \right)^2 \\
&= \sum_{x \in \mathcal{X}} \left(\sum_{j=1}^{\infty} \sum_{k=1}^{j} \sum_{y \in K} \nu_x(j) P^t(x,y) \right)^2 \\
&= \sum_{x \in \mathcal{X}} \big(m(x) P^t(x,K) \big)^2 .
\end{aligned}$$

For b_2, we begin by writing

$$b_2 = \sum_{x \in \mathcal{X}} \sum_{(j,k,y)} \sum_{(j',k',y') \neq (j,k,y)} E\left(H_{(x,j,k,y)} H_{(x,j',k',y')} \right) .$$

We now observe that the expected value will be 0 unless $j = j'$ and $k \neq k'$ (because if $j = j'$ and $k = k'$, then necessarily $y \neq y'$, and a particle cannot be in two different places at once). The sum over all $k \neq k'$ then just contributes a factor $j(j-1)$, and we obtain

$$b_2 = \sum_{x \in \mathcal{X}} \sum_{j=1}^{\infty} \nu_x(j) j(j-1) \sum_{y \in K} \sum_{y' \in K} P^t(x,y) P^t(x,y') = \sum_{x \in \mathcal{X}} (m_2(x) - m(x)) \left(P^t(x,K) \right)^2 .$$

The result now follows. ∎

Remark. The results in [AGG] actually allow for some dependence among the random variables X_α, controlled through the quantity

$$b_3 = \sum_{\alpha} E \left| E \left(X_\alpha - E(X_\alpha) \mid \sum_{\beta \notin B_\alpha} X_\beta \right) \right| .$$

It is possible that this more general bound could be used to obtain results which do not require the assumption that our initial distribution ν be a product measure.

5. Rates for specific examples.

It is natural to ask what our results say about specific examples. We begin with a careful consideration of simple symmetric random walk. We then provide an example to demonstrate that it is possible for our system chains to converge exponentially quickly.

Proposition 7. *Let $\{N_x(t)\}_{x\in\mathbf{Z},\, t\in\mathbf{Z}_+}$ represent the system chain corresponding to simple symmetric random walk on the integers, started with exactly one particle at each site. Assume for simplicity that t is even. Then*

$$(3/4e)\binom{2t}{t}/4^t \;\leq\; \left\| \mathcal{L}\left(\{N_x(t)\}_{x\in K}\right) \;-\; \prod_{x\in K} Poisson\,(m(x)) \right\|_{\mathrm{var}} \leq\; 4\,|K|^2\,\binom{2t}{t}/4^t\,.$$

Proof. This follows from Theorem 4 and the remark following. Note that here $m(x) = h_\gamma = F_K = 1$, and for any $x, y \in \mathbf{Z}$, $P^{2t}(x,y) \leq P^{2t}(x,x) = \binom{2t}{t}/4^t$. The upper bound follows immediately. The lower bound follows from this and from recalling again that $1 - e^{-r} \geq 3r/4$ for $r \leq \frac{1}{2}$. ∎

By Stirling's approximation, $\binom{2t}{t}/4^t = (1 + o(1))/\sqrt{\pi t}$ as $t \to \infty$. Thus, the above proposition shows that the system chain corresponding to simple symmetric random walk converges at rate $1/\sqrt{t}$.

We now turn our attention to the question of exponentially fast convergence of our system chain. Corollary 5 shows that if the initial distribution satisfies (A1), then this convergence will be exponential if and only if the quantity $s_K(t)$ decreases exponentially quickly. Assuming there is at least one particle created at each site of $\mathcal{X}$, with probabilities bounded away from 0, this is equivalent to saying that $\sup\limits_{x\in\mathcal{X}} P^t(x,y)$ decreases exponentially quickly (in t) for each $y \in K$.

We now observe that under these circumstances, if $P^t(x,y)$ is *recurrent*, then the convergence of the system chain will be sub-exponential. Indeed, if it were exponential, then $P^t(x,x)$ would decrease exponentially quickly for any $x \in K$, and so $\sum\limits_{t=1}^{\infty} P^t(x,x)$ would be finite, implying that $P^t(x,y)$ is transient.

Similar comments apply if $P^t(x,y)$ satisfies a weaker condition than recurrence, namely that for a given $y \in K$ there are deterministic points $x_1, x_2, \ldots$, such that $\sum_{t=1}^{\infty} P^t(x_t, y)$ is infinite. (Recurrence is equivalent to being able to take $x_t = y$ for all t.)

However, there are indeed well-behaved transient walks for which the convergence is exponential. We give one example here.

Proposition 8. *Let $P^t(x,y)$ correspond to simple random walk on the doubly-infinite binary tree. (Thus, at each step a particle goes to one of the 3 neighboring sites with probability 1/3.) Then, there is $a > 0$ such that for any sites x and y,*

$$P^t(x,y) < e^{-at}$$

for sufficiently large t. Thus, the corresponding system chain, beginning with 1 particle at each site, converges exponentially quickly.

Proof. Note that at each step, a particle's distance to x will increase with probability 2/3, and decrease with probability 1/3 (assuming the particle is not exactly at x). Thus, by the large deviation principle, for $\epsilon > 0$ the probability of a particle remaining within $(\frac{1}{3} - \epsilon)t$ of x after time t is bounded by e^{-rt} for some $r > 0$. Also, if the particle is at a point y which is a distance $j > (\frac{1}{3} - \epsilon)t$ away from x, it is equally likely to be at any of the 2^j equidistant points, so $P^t(x,y) \leq 2^{-j}$. Combining these facts, the result follows. ∎

REFERENCES

[AGG] R. Arratia, L. Goldstein, and L. Gordon (1989), *Two moments suffice for Poisson approximations: the Chen-Stein method.* Ann. Prob. **17**, 9-25.

[AT] R. Arratia and S. Tavaré (1993), Book review. Ann. Prob. **21**, 2269-2279.

[BHJ] A.D. Barbour, L. Holst, and S. Janson (1992), *Poisson approximation.* Oxford University Press.

[Ch] L. H. Y. Chen (1975), *Poisson approximation for dependent trials.* Ann. Prob. **3**, 534-545.

[Di] P. Diaconis (1988), *Group Representations in Probability and Statistics,* IMS Lecture Series vol. **11**, Hayward, CA.

[DS] P. Diaconis and D. Stroock (1991), *Geometric bounds for eigenvalues of Markov chains,* Ann. Appl. Prob. **1**, No. **1**, 36-61.

[De] J.-D. Deuschel (1994), *Algebraic L^2 decay of attractive processes.* Ann. Prob. **22**, No. **1**, 264-283.

[Dob] R.L. Dobrushin (1956), *On Poisson laws for distributions of particles in space.* Ukrain. Math. Z. **8**, 127-134. (In Russian.)

[Doo] J.I. Doob (1953), *Stochastic Processes,* Wiley, New York. Chapter $VIII$, Section 5.

[Dy] E.B. Dynkin (1983), *Markov Processes as a tool in field theory,* J. Func. Anal. **50**, 167-187.

[Fe] W. Feller (1953), *Introduction to Probability Theory and its Applications, Vol. I,* Wiley, New York.

[Li] T.M. Liggett (1978), *Random invariant measures for Markov chains, and independent particle systems.* Z. Warhschemlichkeitstheorie verw. Gebiete **45**, 297-313.

[LP] T.M. Liggett and S.C. Port (1988), *Systems of independent Markov chains.* Stoch. Proc. Appl. **28**, 1-22.

[Ste] C. Stein (1971), *A bound for the error in the normal approximation to the distribution of a sum of dependent random variables.* In *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **3,** 583-602. University of California Press.

[Sto] C. Stone (1968), *On a theorem by Dobrushin,* Ann. Math. Stat. **39**, 1391-1401.

[R] J.S. Rosenthal (1993), *Minorization conditions and convergence rates for Markov chain Monte Carlo.* Technical Report **9321**, Dept. of Statistics, University of Toronto. J. Amer. Stat. Assoc., to appear.

[VSC] N. Varopoulos, L. Saloff-Coste, and Th. Coulhon (1993), *Analysis and geometry on groups.* Cambridge University Press.