COUNTING VERTICES IN A VOTER-TYPE MODEL VIA STEIN’S METHOD

Radoslav Marinov,* Concordia University College of Alberta

Abstract

The Neighborhood Attack model is a Voter type model, which takes a finite graph, assigns 1’s and -1’s to its nodes (vertices), and then runs a Markov chain on the graph by uniformly at random picking a node at every turn, and then switching the values of the node and its neighbors to 1’s or -1’s according to a (not necessarily fair) coin toss. We show, via a Stein’s method argument, that for certain (highly symmetric) families of graphs the number of 1’s in the Neighbourhood Attack Voter-type model is asymptotically normally distributed as the number of nodes tends to infinity.

Keywords: Stein’s method; Markov chains; Voter models; Neighborhood Attack model; interacting particle systems; bounds of convergence

2010 Mathematics Subject Classification: Primary 05C81

Secondary 60J05; 82B20

1. Introduction and background

In this paper, we seek to apply Stein’s method – a technique for obtaining convergence (often CLT-type) results for random variables – on a vertex-count in the Neighborhood Attack Voter-type model.

Voter models are interacting-particle-system models on finite graphs. The original Voter model (introduced independently in the 1970s by Clifford and Sudbury in 1973, and by Holley and Liggett in 1975, as mentioned in [13]) can be formulated as follows: Take a connected, r-regular (each vertex has r edges) graph of size n. Assign 1’s and -1’s to the nodes of the graph. Run a Markov chain on the graph with the following transition procedure: each turn, pick a node at random (under some distribution;
usually we take the uniform), pick one of its neighbors at random (usually uniformly), and switch the value of the selected neighbor-node to the value of the originally selected node. Under uniformity of node and neighbor selection, this chain converges to one of two absorbing states, in which all nodes have the same values.

The “Anti-voter” model, introduced in [14], has the selected neighbour node adopt a value opposite to that of the originally selected node. Under uniformity (again, of node and neighbor selection) the resulting chain has a stationary distribution.

Persi Diaconis and Christos Athanasiadis in [2] proposed the following variation of the Voter model: upon selecting a node, instead of picking one of its neighbors, flip a coin (with weight \( p \), perhaps taken to be a half), and, according to the result of the cointoss, assign either 1 or \(-1\) to the selected nodes and all its neighbors. The model has been labeled the “Neighborhood Attack” model.

Stein’s method (first introduced in [19]) provides an infrastructure for the estimation of the distances between certain classes of random variables and certain (usually classical) distributions, most notably the Gaussian and the Poisson distributions. For practical purposes, we can break Stein’s method into three key steps: First, one has to use Stein’s identities to establish a bound on the distance between a class of random variables and a specific distribution expected to be close to the given class; second, one has to satisfy the conditions generated in the preceding step; and third, one has to evaluate the acquired bound. The last step typically involves something along the lines of reducing an expression involving a function of the variance of the given random variable.

In [16], Yosef Rinott and Vladimir Rotar show, using a Stein’s method argument, that the sum of the values of the nodes in the Anti-voter model at stationarity is asymptotically normally distributed. The problem Rinnott and Rotar tackled was posed by Aldous and Fill in a book that touches on Voter models, [1]. Our goal in the present article is to show that the sum of the values of the nodes in the Neighborhood Attack model is asymptotically normally distributed, using Stein’s method techniques different from the ones employed by Rinott and Rotar.

For an application of the Stein technique in a different context, see the paper [9], in which Jason Fulman shows that the number of descents or inversions in permutations complies to a central limit theorem. Both the current problem and the one examined
counting vertices in a voter-type model

in [9] can be viewed as random walks on hyperplanes; and hence there is a structural
similarity between the approach adopted here, and the one in [9].

for more results on the neighbourhood attack model, see [2], [6]. the former
paper introduces the model and presents some results on random walks on hyperplane
arrangements. the latter paper studies some properties of the distributions of the
implicit markov chains in models similar to the neighbourhood attack model.

for more on stein’s method, see [5], [3], [18]. the first two books provide a
comprehensive overview of stein’s method in regard to its applications to normal
and poisson approximations reflexively. the monograph [18] is an up-to-date survey
of stein’s method literature and a useful entry-level source on the subject.

in section 2, we pose our problem. in section 3, we conduct a brief overview of
our main technique: stein’s method. in section 4, we introduce a few definitions
and assumptions, and then list the main result of the paper. in the section 5, we
provide calculations and proofs for the result. section 6 interprets the result with
some examples of its applicability. we draw conclusions in section 7.

2. problem and approach

we apply the neighbourhood attack model (introduced in [2]) on a given family
of (finite) graphs. randomly assign either 1 or −1 to each node of the graph. as
mentioned above, the model does the following each turn:

- selects a node uniformly at random.
- turns the node and all its immediate neighbours into 1’s or −1’s according to
  a bernoulli(p) distribution with 0 < p < 1; we want p = 1/2 for the sake of
  symmetry.

given: 1) a connected graph; 2) positive probability of selection for all nodes; and 3)
positive probabilities of turning into 1 or −1 for the selected node and its neighbours,
the underlying markov chain, the states of which are the possible permutations of 1’s
and −1’s, is irreducible and everywhere recurrent on an essential class of its state space,
and therefore possesses a stationary distribution. assume the considered markov chain
begins at this stationary distribution.
Let $X$ be the number of 1’s at stationarity. Then $N - X$ equals the number of −1’s, where $N$ is the number of nodes.

We want to use Stein’s method to show that

$$\frac{X - \mathbb{E}X}{\sigma_X} \xrightarrow{N \to \infty} Z$$

where $Z$ is of the standard normal distribution, and $\mathbb{E}X$ and $\sigma_X$ are the expectation and standard deviation of $X$.

We derive our result under an assumption of $r$-regularity for the underlying graphs.

We seek to apply Stein’s method, and in particular we want to use a result along the lines of Theorem 1.2 in [16]:

**Theorem 1.** Let $(W, W')$ be exchangeable with $\mathbb{E}W = 0$ and $\mathbb{E}W^2 = 1$. Define the r.v. $R = R(W)$ by

$$\mathbb{E}(W'|W) = (1 - \lambda)W + R,$$

where $0 < \lambda < 1$. Then, if there is some $A$ for which $|W' - W| \leq A$, we have

$$\delta := \sup \{ |\mathbb{E}h(W) - \Phi h| : h \in \mathcal{H} \} \leq$$

$$\leq \frac{12}{\lambda} \sqrt{\text{Var} \{ \mathbb{E}[(W' - W)^2 | W] \}} + \frac{37 \sqrt{\mathbb{E}R^2}}{\lambda} + 48 \frac{aA^3}{\lambda} + 8 \frac{aA^2}{\sqrt{\lambda}},$$

where $\mathcal{H}$ is such that all functions in it are uniformly bounded in absolute value by 1, for any real numbers $c$ and $d$ and any $h \in \mathcal{H}$, the function $h(cx + d)$ is in $\mathcal{H}$, and for any $\epsilon > 0$ and any $h \in \mathcal{H}$, the functions $h^+_\epsilon, h^-_\epsilon$ are also in $\mathcal{H}$, and

$$\int h(x; \epsilon)\Phi(dx) \leq a\epsilon$$

for some constant $a$ which depends only on the class $\mathcal{H}$.

Our $W$ would be some normalization of a vertex-count on the Voter-type model graphs we deal with.

3. Brief overview of Stein’s method

Stein’s technique goes as follows: for a given probability distribution, one can come up with an appropriate operator which implicitly defines the distribution. For example, the operator $A$ in $Af(x) = f'(x) - xf(x)$ implicitly defines the Gaussian distribution,
in the sense that 1) $\mathbb{E} Af(Z) = 0$ for all absolutely continuous $f$ with $\mathbb{E}|f'(Z)| < \infty$, where $Z$ is a variable with the standard normal distribution; and 2) if for some random variable $W$ we have $\mathbb{E} Af(W) = 0$ for all absolutely continuous functions $f$ with $|f'| < \infty$, then $W$ has the standard normal distribution.

Next, for an appropriately chosen $A$, one can solve the differential equation given by

$$Af(x) = 1_{w \leq x} - \Phi(x), \tag{2}$$

where $\Phi(x)$ is the c.d.f. of the target distribution.

But now, armed with the solution to equation (2), and within the context of an appropriate metric (above we used the Kolomogorov metric), we can produce a bound on the distance $|P(W \leq x) - \Phi(x)|$ between a given distribution we want to analyze, and the target distribution with c.d.f. $\Phi(x)$.

For example,

$$f_x(w) = e^{w^2/2} \int_w^\infty e^{-t^2/2} (\Phi(x) - 1_{t \leq x}) \, dt$$

is the unique bounded solution to

$$f'_x(w) - w f_x(w) = 1_{w \leq x} - \Phi(x),$$

where $\Phi(x)$ is the c.d.f. of the standard normal. And next, under the Wasserstein metric given by $\mathcal{H} = \{ h : \mathbb{R} \rightarrow \mathbb{R} : |h(x) - h(y)| \leq |x - y| \}$, one can show that (for example, see [18, 3.1])

$$d_W(W, Z) \leq (A + B)n^{-1/2}, \quad A = \mathbb{E}|X_1|^3, \quad B = \frac{\sqrt{2\mathbb{E}|X_1^4|}}{\sqrt{\pi}},$$

where $W$ is a normalized sum of $n$ i.i.d. standard normal variables endowed with a fourth moment, and $d_W(W, Z)$ stands for the Wasserstein distance between $W$ and the standard normal distribution.

The potential utility of Stein’s technique in producing powerful bounds and obtaining convergence results is clear; and, indeed, Stein’s method has been instrumental in the proofs of a variety of interesting convergence and bounding results. In general, there are two standard avenues of research focusing on Stein’s method – one can try to obtain formulas for bounds on the distances between various target distributions and various random variables (or rather, their distributions) – examples of recent results
in this direction include [10] (Exponential distribution), [15] (Laplace), and [11] (zero-bias couplings and concentration inequalities); and one can use these formulas and techniques to obtain results pertaining to specific problems, including many classic problems such as the Birthday Problem or the Coupon Collector Problem – for examples, refer to [4] (comprehensive survey) and [12] (Lightbulb process).

4. Initial setup and main result

4.1. Initial setup

We first seek to show that (1) holds. To that end, let \( X \) be the number of 1’s at stationarity. Let

\[
Y = 2X - N = \sum_{i=1}^{N} \xi_i.
\]

Here \( N \) is the total number of nodes and \( \xi_i \) is the value of node \( i \) (under an arbitrary indexing). Examining \( Y \) is equivalent to examining \( X \). Next, define

\[
W := \frac{Y - \mathbb{E}Y}{\sigma_Y}.
\]

Note \( \sigma_Y \) is a constant dependent on \( N \):

\[
\sigma_Y^2 = \text{Var} \left( \sum_{i=1}^{N} \xi_i \right) = \sum_{i=1}^{N} \text{Var} \xi_i + 2 \sum_{1 \leq i < j \leq N} \text{Cov}(\xi_i, \xi_j)
\]

Now, \( W \) is mean-0 variance-1. To get the condition for Theorem 1, we first need to define a \( W' \) as the equivalent of \( W \) after one further turn of the Neighborhood Attack model. That is to say, if \( W \) is the normalized node count of the model at some turn of its evolution in stationarity, then \( W' \) is the same normalized node count in the next turn.

Note, once again, that we assume \( r \)-regularity for the graph (i.e. every node has exactly \( r \) neighbours). We want \( r \)-regularity for the sake of symmetry, because without symmetry, the problem under consideration is far less tractable.

4.2. Main result

**Theorem 2.** Under the assumptions

\[
r_i^* = r^* \quad \forall i \in I,
\]

(3)
where $I$ is the index-set of the nodes, $r^*_i$ is the number of first or second order neighbors node $i$ has, and $r^*$ is some constant dependent on the graph; and

$$\text{Cov}(\eta, \theta) \leq 0,$$

(4)

where $\eta$ is the count of pairs of neighbors or near-neighbors with values both equal to 1, and $\theta$ is the count of pairs of $-1$’s, we derive the bound on the distance between (the distributions of) $W$ and the standard normal,

$$\delta \leq 48\frac{r^2}{\sqrt{r + 1}\sqrt{N}} + (2^{19/2} + 48)\frac{\sqrt{r + 1}}{\sqrt{N}}.$$

(5)

We first establish bounds on $\sigma_Y$ in Section 5.3 and then complete the proof in Section 5.4.

5. Details and proof

5.1. Proving $\mathbb{E}(W' | W) = (1 - \lambda)W$

Given the $r$-regularity assumption, the sum $\left(\sum_{i=1}^{N} \xi_i\right)$ changes each turn by between $-2(r+1)$ and $2(r+1)$. A basic example of a graph of this type is the circle (2-regular) graph, in which we have a set of nodes arranged in a circle, each node with two neighbours.

We also assumed uniformity in choosing nodes and in flipping 1’s or $-1$’s. Under such conditions $\mathbb{E}\xi_i = 0$, i.e. at stationarity each node is 1 or $-1$ with equal probability. The sum of the node values will tend toward 0 (under certain conditions; one of which, clearly, has to do with the number of neighbors each node has, since our model takes only the extreme values over the complete graph), since if nodes of a certain value ($+1$ or $-1$) dominate the graph, we are less likely to see an increase in the number of the nodes of that value.

We show in Section 5.3 that, as desired,

$$\mathbb{E}(W' | W) = \left(1 - \frac{r + 1}{N}\right)W,$$

which complies with the Stein linearity condition

$$\mathbb{E}(W' | W) = (1 - \lambda)W + R,$$
where $0 \leq \lambda \leq 1$ and $R$ is a random variable. In our case conveniently $R = 0$. As for lambda,
\[
\lambda = \frac{r + 1}{N}.
\]

### 5.2. Röllin’s result

In general, the next step is to show that $W$ and $W'$ are exchangeable, i.e. $(W, W') =^d (W', W)$, as was done in [9]. Exchangeability clearly holds when the Markov chain underlying $W$ and $W'$ is stationary and reversible. Reversibility is not always available or easily proved.

For example, our chain is clearly not necessarily reversible. Consider the circle graph. It is easy to see that for $N$ large, $Y$ can take the value $N - 2$ — i.e. there is an attainable at stationarity arrangement of values for the nodes in which all nodes but one have the value of 1. Now, the probability of going from that arrangement to the all 1’s arrangement for which $Y = N$ is positive; but the probability of going from $Y = N$ to $Y = N - 2$ is zero, and hence our chain fails to satisfy the detailed balance equations $\pi(x)P(x, y) = \pi(y)P(y, x)$.

However, a recent result by Adrian Röllin removes the necessity for exchangeability. Röllin’s theorem (see [17, Theorem 2.1]) states:

**Theorem 3.** Assume $W, W'$ are r.v.s on the same probability space, s.t. $\mathcal{L}(W') = \mathcal{L}(W)$ ($\mathcal{L}$ for ‘law’), $E W = 0$, $\text{Var}(W) = 1$. Given $E(W'|W) = (1 - \lambda)W + R$, for
\[
\delta := \sup_{h \in \mathcal{H}} |E h(W) - E h(Z)|
\]
where $Z$ is the standard normal distribution, and $\mathcal{H}$ is the family of functions associated with the Wasserstein distance, we have
\[
\delta \leq \frac{6}{\lambda} \sqrt{\text{Var}(W)(W' - W)^2} + \frac{19\sqrt{\text{Var}(R)}}{\lambda} + 4\sqrt{\frac{a\mathbb{E}|W' - W|^3}{\lambda}}.
\]

If also there exists a constant $A$ s.t. $|W' - W| \leq A$ a.s., we have
\[
\delta \leq \frac{12}{\lambda} \sqrt{\text{Var}(W)(W' - W)^2} + \frac{37\sqrt{\text{Var}(R)^2}}{\lambda} + \frac{32}{\lambda} A + 6 \frac{A^2}{\sqrt{\lambda}}.
\]

**Proof.** See [17, Theorem 2.1].

In our case $R = 0$; so the bound is
\[
\delta \leq \frac{12}{\lambda} \sqrt{\text{Var}(W)(W' - W)^2} + \frac{32}{\lambda} A + 6 \frac{A^2}{\sqrt{\lambda}}. \quad (6)
\]
The next step is to bound $|W' - W|$. Note $|Y' - Y| \leq 2(r + 1)$. So

$$|W' - W| \leq \frac{2(r + 1)}{\sigma_Y} = A.$$ 

Thus, the bound becomes

$$\delta \leq \frac{12N}{(r + 1)\sigma_Y^2} \sqrt{\text{Var}[E(Y' - Y)^2|Y]} + 32\frac{(r + 1)^2N}{\sigma_Y^3} + 6\frac{4(r + 1)^{3/2}\sqrt{N}}{\sigma_Y^2}. \quad (7)$$

\[5.3.\text{ Bounding the variance of}\ Y\]

In effect, the next goal is to bound the two terms $\text{Var}[E(Y' - Y)^2|Y]$ and $\text{Var}(Y') = \text{Var}(Y) = \sigma_Y^2$. For an explicit formula for the value of the corresponding $\sigma_Y$ in the Anti-voter model, see [1, Chapter 14]. Another relevant paper dealing with the Anti-voter case can be found in [7].

Let us first try to find $\text{Var}_Y$. Define $\Delta Y = Y' - Y$ to obtain $Y' = Y + \Delta Y$.

By stationarity, it follows that $0 = \text{Var}(Y') - \text{Var}(Y)$. Hence:

$$0 = E(Y')^2 - EY^2 =$$

$$= E \left[ Y^2 + 2Y\Delta Y + (\Delta Y)^2 - Y^2 \right] =$$

$$= 2E(Y\Delta Y) + E(\Delta Y)^2. \quad (8)$$

To continue, we need to obtain a firmer grip on the r.v. $\Delta Y$. It is easy to see that $\Delta Y$ takes values between $-2(r + 1)$ and $2(r + 1)$, and that the probability distribution of $\Delta Y$ is a function of certain edge and vertex counts on the coloured graph, themselves random variables. Specifically:

Define $q_i$ as the number of nodes s.t. the sum of the values at the node and all its neighbors equals $i$. Clearly, $i$ takes integer values (all odd or all even depending on the parity of $r$) between $-(r + 1)$ and $(r + 1)$. Specifically, if $r$ is odd, $i$ takes the values $-(r + 1), -(r - 1), ..., -2, 0, 2, ..., (r - 1), (r + 1)$; and if $r$ is even, $i$ takes the values $-(r + 1), -(r - 1), ..., -1, 1, ..., (r - 1), (r + 1)$. In each case $i$ takes $(r + 2)$ distinct values. Call the set of those values (the possible values $i$ can take) $I$.

Simple counting produces two useful identities involving the $q_i$’s:

$$\sum_{i \in I} q_i = N \quad \text{and} \quad \sum_{i \in I} iq_i = (r + 1)Y.$$ 

Now, at each turn of the Neighborhood Attack process we pick a node uniformly at random (i.e. with probability $1/N$), and turn its value and the value of all its neighbors to either 1 or $-1$ uniformly at random (i.e. with probability $1/2$).
It thus follows that $\Delta Y$ has the (conditional on $\{q_i\}$) p.d.f.:

$$
\Delta Y = \begin{cases} 
(r + 1) - i & \text{with probability } q_i \frac{1}{2N} \\
-(r + 1) - i & \text{with probability } q_i \frac{1}{2N}
\end{cases}
$$

So for example, $\Delta Y$ takes the value of $2(r + 1) = (r + 1) - (-r + 1)$ with probability $q_{r+1}/2N$; and the value $-2(r + 1) = -(r + 1) - (r + 1)$ with probability $q_{-(r+1)}/2N$.

Thus, we have

$$
E[\Delta Y | Y] = EE[(\Delta Y | \{q_i\}) | Y] =
$$

$$
= E \left( \sum_{i \in I} [(r + 1) - i] \frac{q_i}{2N} + \sum_{i \in I} [-(r + 1) - i] \frac{q_i}{2N} | Y \right) =
$$

$$
= -2E \left( \sum_{i \in I} \frac{q_i}{2N} | Y \right) = -\frac{(r + 1)Y}{N},
$$

which is what we stated in Section 5.1.

So:

**Lemma 1.** In the Neighborhood Attack model on an $r$-regular graph,

$$
E[\Delta Y | Y] = -\frac{(r + 1)Y}{N}.
$$

That is, the Stein linearity condition is satisfied with $\lambda = \frac{(r+1)}{N}$ and $R = 0$.

We continue from (8):

$$
0 = E(Y')^2 - EY^2 =
$$

$$
= 2E(Y \Delta Y) + E(\Delta Y)^2 =
$$

$$
= -\frac{2(r + 1)}{N} E(Y^2) + E \left( \sum_{i \in I} [(r + 1 - i)^2 + (-(r + 1) - i)^2] \frac{q_i}{2N} | \{q_i\} \right)
$$

(9)

Let us focus on the $E \left( \sum_{i \in I} [(r + 1 - i)^2 + (-(r + 1) - i)^2] \frac{q_i}{2N} | \{q_i\} \right)$ term:

$$
E \left( \sum_{i \in I} [(r + 1 - i)^2 + (-(r + 1) - i)^2] \frac{q_i}{2N} | \{q_i\} \right) =
$$

$$
= (r + 1)^2 + \frac{1}{N} \left( \sum_{i \in I} i^2 q_i \right) \leq
$$

$$
\leq (r + 1)^2 + \frac{1}{N} \left( \sum_{i \in I} (r + 1)^2 q_i \right) = (r + 1)^2 + (r + 1)^2 = 2(r + 1)^2
$$

(10)
And therefore, continuing from (9),

\[ 0 = \mathbb{E}(Y')^2 - \mathbb{E}Y^2 = \]

\[ = \frac{-2(r + 1)}{N} \mathbb{E}(Y^2) + \mathbb{E} \left( \sum_{i \in I} \left[ (r + 1 - i)^2 + (- (r + 1) - i)^2 \right] \frac{q_i}{2N} \{\{q_i\} \right) \leq \]

\[ \leq \frac{-2(r + 1)}{N} \mathbb{E}(Y^2) + 2(r + 1)^2, \]

meaning

\[ \sigma_Y^2 = \text{Var}(Y) = \mathbb{E}(Y^2) \leq (r + 1)N. \]

However, since the \( \sigma_Y \) terms appear in the denominators of the terms in (7), we need either a lower bound of \( \sigma_Y \) or the exact variance of \( Y \).

Observe that we have:

\[ \text{Var}(Y) = \mathbb{E}Y^2 = \frac{N}{2(r + 1)} \mathbb{E} \left( \sum_{i \in I} \left[ (r + 1 - i)^2 + (- (r + 1) - i)^2 \right] \frac{q_i}{2N} \{\{q_i\} \right) = \]

\[ = \frac{N}{2(r + 1)} \left[ (r + 1)^2 + \frac{1}{N} \mathbb{E} \left( \sum_{i \in I} i^2 q_i \right) \right] \geq \frac{(r + 1)N}{2}. \]

Thus

**Lemma 2.** For the Neighborhood Attack model on an \( r \)-regular graph, for \( Y = \sum_{i=1}^{N} \xi_i \) the sum of the values of the nodes of the graph,

\[ \frac{(r + 1)N}{2} \leq \sigma_Y^2 \leq (r + 1)N. \]

5.4. Reducing and bounding \( \text{Var} \mathbb{E}[(Y' - Y)^2 | Y] \)

Now we have to evaluate or bound \( \text{Var} \mathbb{E}[(Y' - Y)^2 | Y] = \text{Var} \mathbb{E}[(Y' - Y)^2] \), as in [16] and [9].

Let us consider the following:

\[ \text{Var} \left[ \mathbb{E}(Y' - Y)^2 | Y \right] = \text{Var} \left[ \mathbb{E}(\Delta Y)^2 | Y \right] \leq \text{Var} \left[ \mathbb{E}(\Delta Y)^2 | \{q_i\} \right] = \]

\[ = \text{Var} \left( (r + 1)^2 + \frac{1}{N} \left[ \sum_{i \in I} i^2 q_i \right] \right) = \frac{1}{N^2} \text{Var} \left( \sum_{i \in I} i^2 q_i \right) \]

The transition between the lines follows from [10] and [9]. Also,

\[ \text{Var} \left[ \mathbb{E}(Y' - Y)^2 | Y \right] = \frac{1}{N^2} \text{Var} \left( \mathbb{E} \left( \sum_{i \in I} i^2 q_i \right) | \{\xi_i\} \right) = \frac{1}{N^2} \text{Var} \left( \frac{1}{N} \left( \sum_{j \in N_k} \xi_j \right) \right), \]
where $N_k$ is the set of node $k$ and all its neighbors.

Next,

$$\operatorname{Var}\left(\sum_{k=1}^{N} \left(\sum_{j \in N_k} \xi_j \right)^2\right) = \operatorname{Var}\left(N(r + 1) + \sum_{0 \leq i, j \leq N; i, j: d(i, j) = 1, 2} \xi_i \xi_j \right) = \operatorname{Var}(2(\alpha - \beta)) = \operatorname{Var}(4(\eta + \theta)).$$

Here $d(i, j)$ is the distance between nodes $i$ and $j$. For the last line, observe that $N(r + 1)$ is invariant, and that the sum $\sum_{0 \leq i, j \leq N; i, j: d(i, j) = 1, 2} \xi_i \xi_j$ can be interpreted as a sort of an edge count over our graph, with each pair of neighbors or near neighbors of the same sign participating as a $+1$, and each pair of opposite values participating as a $-1$. Each such pair gets counted twice.

Thus let $r^*$ be the number of neighbors or near-neighbors (meaning nodes at distances one or two) each node has; $\alpha$ be the number of pairs of neighbors or near-neighbors with equal node-values; and $\beta$ be the number of pairs with opposite node-values. Moreover, $\eta$ corresponds to the count of pairs of neighbors or near-neighbors with values both equal to 1, and $\theta$ is the count of pairs of $-1$’s.

Now, suppose, with $I$ the indexed set of nodes on our graph, that

$$r^*_i = r^* \quad \forall i \in I.$$

That is, we assume $r^*$ is some fixed quantity: i.e. the underlying graph possesses sufficient symmetry so that each node has the same number of neighbors or near-neighbors. For example each node in the circle-graph has 4 neighbors/ near-neighbors.

The assumption that $r^*_i$ is fixed for all $i$ is not particularly gratuitous, since, either way, $r^*_i \leq r^2_i$, and under the present assumptions, $r_i = r$ is fixed.

Next, since $2(\alpha + \beta) = r^*N$, we have $2(\alpha - \beta) = 2\alpha - (r^*N - 2\alpha) = 4\alpha - r^*N$. Therefore, $\operatorname{Var}(2(\alpha - \beta)) = \operatorname{Var}(4\alpha) = \operatorname{Var}(4(\eta + \theta))$, where $\eta$ is the number of pairs of neighbors and near neighbors with node-values 1, and $\theta$ is the count of pairs with values $-1$.

On the other hand, we have $2(\eta - \theta) = r^*Y$, and therefore $\operatorname{Var}(\eta - \theta) = \frac{(r^*)^2}{4} \operatorname{Var}(Y)$. 
One naturally wonders if we can use the bound we established for \( \text{Var}(Y) \) to bound \( \text{Var}(\eta + \theta) \). From the definition,

\[
\text{Var}(\eta + \theta) = \text{Var}(\eta) + 2 \text{Cov}(\eta, \theta) + \text{Var}(\theta)
\]

\[
\text{Var}(\eta - \theta) = \text{Var}(\eta) - 2 \text{Cov}(\eta, \theta) + \text{Var}(\theta)
\]

It suffices to show that \( \text{Cov}(\eta, \theta) \leq 0 \) to obtain \( \text{Var}(\eta + \theta) \leq \text{Var}(\eta - \theta) \). For \( \eta \) and \( \theta \) to be negatively correlated, an increase in one would have to imply a decrease in the other – meaning, in our setting, that an increase in the number of edges (i.e. pairs of nodes at distance 1) with ones at both ends would have to imply a decrease in the edges with negative ones at both ends – and vice versa.

One is tempted to try to use the FKG inequality to prove \( \text{Cov}(\eta, \theta) \leq 0 \). Specifically, we know that the lattice \( \{-1, 1\}^\Gamma \) (where \( \Gamma \) is our graph) is a poset; and that \( \eta = f(\{\xi_i\}) \) is an increasing function of that lattice, while \( \theta = g(\{\xi_i\}) \) is a decreasing function on the same lattice. Moreover, if we take an element \( x \) from \( \{-1, 1\}^\Gamma \), and suppose that \( \overline{x} \) is the element we obtain by switching all \(-1\)'s in \( x \) to \(+1\)'s, and all \(+1\)'s to \(-1\)'s, then \( f(x) = g(\overline{x}) \), and the stationary probability \( p_x \) of state \( x \) occurring in our Markov Chain equals the corresponding probability for state \( \overline{x} \) – that is, \( p_x = p_{\overline{x}} \).

Now, the FKG theorem (after Fortuin, Kasteleyn, and Ginibre [8]) states that for \( X \) a finite distributive lattice, and \( \mu \) a non-negative function (really a measure) on it, satisfying the “log-supermodularity condition”

\[
\mu(x \land y)\mu(x \lor y) \geq \mu(x)\mu(y) \quad \forall x, y \in X
\]

yields

\[
\left( \sum_{x \in X} f(x)g(x)\mu(x) \right) \left( \sum_{x \in X} \mu(x) \right) \geq \left( \sum_{x \in X} f(x)\mu(x) \right) \left( \sum_{x \in X} g(x)\mu(x) \right)
\]

for any two monotonically increasing (or decreasing) on \( X \) functions \( f \) and \( g \); with the inequality reversed if one of \( f \) and \( g \) is monotonically increasing and other one monotonically decreasing.

In our particular case, for the considered lattice, stationary distribution over the lattice, and functions \( f \) and \( g \), having

\[
\left( \sum_{x \in \{-1, 1\}^\Gamma} f(x)g(x)p_x \right) \left( \sum_{x \in \{-1, 1\}^\Gamma} p_x \right) \leq \left( \sum_{x \in \{-1, 1\}^\Gamma} f(x)p_x \right) \left( \sum_{x \in \{-1, 1\}^\Gamma} g(x)p_x \right)
\]
would do the job, since
\[ \mathbb{E}\eta\theta = \sum_{x \in \{-1,1\}^\Gamma} f(x)g(x)p_x, \]
and
\[ \mathbb{E}\eta = \sum_{x \in \{-1,1\}^\Gamma} f(x)p_x = \sum_{x \in \{-1,1\}^\Gamma} g(x)p_x = \mathbb{E}\theta, \]
and
\[ \mathbb{E}\eta\theta \leq (\mathbb{E}\eta)(\mathbb{E}\theta) \]
implies exactly \( \text{Cov}(\eta, \theta) \leq 0. \)

Unfortunately, our example fails to necessarily satisfy the log-supermodularity condition. For example, take a circle graph of odd length. The state in which 1’s and \(-1\)’s alternate along the entire graph cannot occur at stationarity, and therefore has measure zero in the stationary distribution of our chain. But the state \( w_1 \) in which we have one 1 at some node \( i \) and everything else is \(-1\); and the state \( w_2 \) in which node \( i \) and its two neighbors are \(-1\)’s, and the rest of the graph consists of alternating 1’s and \(-1\)’s, can both occur. But then \( p_{w_1}p_{w_2} > 0 \), while \( p_{w_1 \wedge w_2} = 0 \), meaning that the log-supermodularity condition fails. One can come up with similar examples for other standard families of graphs. So we fail to have the log-supermodularity condition.

Still, the log-supermodularity condition is only sufficient rather than necessary for our desired result. Hence our results might be obtainable via different means. For now, suppose
\[ \text{Cov}(\eta, \theta) \leq 0. \]

Given \( \frac{\text{Var} \left( \sum_{i \in I} i^2 q_i \right)}{\text{Var}(2(\alpha - \beta)) = 4 \text{Var}(2(\eta + \theta)) \leq 4 \text{Var}(2(\eta - \theta)) = 4(r^*)^2 \text{Var}(Y) \leq 4(r^*)^2(r + 1)N.} \]

It follows that
\[ \text{Var}[\text{E}(Y' - Y)^2|Y] \leq \frac{4(r^*)^2(r + 1)}{N}. \]

Welsh and Donnelly found that the stationary distribution of the chain underlying the Anti-voted model also fails to satisfy the log-superlinearity condition – see [7, Section 5].
We thus arrive at the overall bound (7):

\[ \delta \leq \frac{12N}{(r+1)\sigma_Y^2} \sqrt{\text{Var}\{ (Y' - Y)^2 \vert Y \}} + 32\frac{8(r+1)^2N}{\sigma_Y^4} + 6\frac{4(r+1)^{3/2}\sqrt{N}}{\sigma_Y^2} \leq \]

\[ \leq 48\frac{r^*}{\sqrt{r+1}\sqrt{N}} + 2^{19/2}\frac{\sqrt{r+1}}{\sqrt{N}} + 48\frac{\sqrt{r+1}}{\sqrt{N}} \]

Thus our overall bound is:

\[ \delta \leq 48\frac{r^*}{\sqrt{r+1}\sqrt{N}} + (2^{19/2} + 48)\frac{\sqrt{r+1}}{\sqrt{N}}, \quad (12) \]

where \( r^* \) is a constant dependent on the underlying family of graphs and satisfying \( r^* \leq r^2 \).

This completes the proof of Theorem 2 and derives (5). The final bound is of \( O\left( \frac{r^*}{\sqrt{r}} N^{-1/2} \right) \).

6. Consequences and explanation of main result

The bound in (12) implies that (under stationarity) the normalized sum of values of the nodes of the graph, \( W \), goes in law to the standard normal distribution as the size of the graph rises given \( \frac{(r^*)^2}{rN} \rightarrow 0 \). Note that \( r \leq r^* \leq r^2 \).

Let us consider four specific families of graphs.

First, the complete graph, in which \( r = N - 1 \). On the complete graph, \( Y = \sigma_Y W \) clearly has the uniform binary distribution taking values \( \pm N \). Thus it is to no surprise that our bound on the distance to the normal distribution rises to infinity with \( N \).

From the other side of the spectrum of regular graphs, we can take the circuit (or circle or simple cycle) graph, in which we have \( N \) ordered nodes, each connected to its predecessor and its successor, with node \( N \) connected to nodes \( N-1 \) and 1. Here \( r = 2 \), and hence \( r^{3/2}/N^{1/2} \) goes to 0 as \( N \) increases to infinity.

The argument can be extended to circulant graphs: as long as \( r \) stays constant as \( N \) rises, \( Y \) would converge to the normal in distribution.

---

A circulant graph is such that we can arbitrarily index its nodes with \( 0,1,\ldots,N-1 \), in such a way that if the nodes corresponding to two indices \( x \) and \( y \) are adjacent, then any two nodes indexed by \( z \) and \( (z - x + y) \mod N \) are adjacent. Here \( N \) is the number of nodes and adjacency of two nodes means they are connected by an undirected edge.
For a slightly more complicated example, consider the hypercube graph. One can index the nodes of the \(n\)-dimensional hypercube graph with a string of \(n\) zeros and ones, with nodes differing in exactly one digit being neighbors.

It is easy to see that for an \(n\)-dimensional hypercube, \(r = n\), and \(N = 2^n\). Since

\[
\frac{r^{3/2}}{N^{1/2}} = \frac{n^{3/2}}{2n^{1/2}} \xrightarrow{n \to \infty} 0,
\]

we can conclude that \(W\) goes in law to the standard normal distribution for the hypercube family of graphs.

Finally, consider the (complete) bipartite graph of size \(N = 2M\), with \(M\) a natural number. For this family, \(r = M\), and \(N = 2M\). On such a graph, \(Y\) would frequently take values near \(-N, 0\) and \(N\), and hence cannot be expected to go to the normal in distribution. Indeed, we have

\[
\frac{r^{3/2}}{N^{1/2}} = \frac{M^{3/2}}{(2M)^{1/2}} \xrightarrow{M \to \infty} \infty.
\]

The argument can clearly extend to multipartite graphs of a fixed number of partitions.

7. Conclusions

To sum up, we have shown that, subject to some symmetry assumptions, the normalized sum of the values of the nodes in the Neighborhood Attack model is at a distance of \(O\left(\frac{r^{3/2}N^{-1/2}}{N^{1/2}}\right)\) to the standard normal distribution in the Wasserstein metric. Hence the sum of the nodes is asymptotically normally distributed as the sizes of the underlying graphs increase, provided that \(r^{3/2}N^{-1/2}\) goes to zero as \(N\) rises to infinity.

Along the way to the result, we also showed that the node-sum \(Y\) in the Neighborhood Attack model on an \(r\)-regular graph satisfies Stein’s linearity condition with \(\lambda = \frac{r+1}{N}\) and \(R = 0\); and that \(\sigma_Y\) satisfies \(\frac{(r+1)N}{2} \leq \sigma_Y^2 \leq (r+1)N\).

Acknowledgement

This work is based on the author’s 2008-2013 graduate research at the University of Southern California under the advisorship of Prof. Jason Fulman. The author would like to express his gratitude to Prof. Fulman and the USC Department of Mathematics.
References

[1] ALDOUS, D. AND A., F. J. (1994-2002). Reversible Markov Chains and Random Walks on Graphs. http://www.stat.berkeley.edu/aldous/RWG/book.html.

[2] ATHANASIADIS, C. A. AND DIACONIS, P. (2010). Functions of random walks on hyperplane arrangements. Advances in Applied Mathematics 45, 410–437.

[3] BARBOUR, A., HOLST, L. AND JANSON, S. (1992). Poisson Approximation vol. 2 of Oxford Studies in Probability. Clarendon Press, Oxford.

[4] CHATTERJEE, S., DIACONIS, P. AND MECKES, E. (2005). Exchangeable pairs and Poisson approximation. Probability Surveys 2, 64–106.

[5] CHEN, L. H. Y. AND GOLDSTEIN, LARRY; SHAO, Q.-M. (2011). Normal Approximation by Stein’s Method. Probability and its Applications. Springer-Verlag Berlin Heidelberg.

[6] CHUNG, F. AND GRAHAM, R. (2012). Flipping edges and vertices in graphs. Advances in Applied Mathematics 48, 37–63.

[7] DONNELLY, P. AND WELSH, D. (1984). The antivoter problem: Random 2-colourings of graphs. Graph Theory and Combinatorics.

[8] FORTUIN, C. M., KASTELEYN, P. W. AND GINibre, J. (1971). Correlation inequalities on some partially ordered sets. Communications in Mathematical Physics 22.

[9] FULMAN, J. (2004). Stein’s Method and non-reversible Markov chains. Lecture Notes - Monograph Series 46, 66–74.

[10] FULMAN, J. AND ROSS, N. (2013). Exponential approximation and Stein’s Method of exchangeable pairs. ALEA, Lat. Am. J. Probab. Math. Stat. 10, 1–13. Arxiv: http://arxiv.org/abs/1207.5073.

[11] GOLDSTEIN, L. AND ISLAK, U. (2013). Concentration inequalities via zero bias couplings. Arxiv: http://arxiv.org/abs/1304.5001.
[12] Goldstein, L. and Zhang, H. (2011). A Berry-Esseen theorem for the lightbulb process. *Advances in Applied Probability* **43.3**, Arxiv: http://arxiv.org/abs/1001.0612.

[13] Liggett, T. M. (2005 reprint of 1985 original publication). *Interacting Particle Systems* vol. 276 of *Classics in Mathematics*. Springer Berlin Heidelberg.

[14] Matloff, N. S. (1977). Ergodicity conditions for a dissonant voting model. *Ann. Probab.* **5**, 371–386.

[15] Pike, J. and Ren, H. (2012). Stein’s method and the Laplace distribution. Arxiv: http://arxiv.org/abs/1210.5775.

[16] Rinott, Y. and Rotar, V. (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted u-statistics. *The Annals of Applied Probability* **7**, 1080–1105.

[17] Röllin, A. (2008). A note on the exchangeability condition in Stein’s Method. *Statistics & Probability Letters* **78**, 1800–1806.

[18] Ross, N. (2011). Fundamentals of Stein’s Method. *Probability Surveys* **8**, 210–293.

[19] Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proc. Sixth Berkeley Symp. on Math. Statist. and Prob.* **2**, 583–602.