Title: *A* Nonlinear Transform for the Diagonalization of the Bernoulli-Laplace Diffusion Model and Orthogonal Polynomials

By *Chjan* C Lim and William Pickering

Department of Mathematical Sciences, Rensselaer Polytechnic Institute

Abstract. The Bernoulli-Laplace model describes a diffusion process of two types of particles between two urns. To analyze the finite-size dynamics of this process, and for other constructive results we diagonalize the corresponding transition matrix and calculate explicitly closed-form expressions for all eigenvalues and eigenvectors of the Markov transition matrix T_{BL}. This is done by a new method based on mapping the eigenproblem for T_{BL} to the associated problem for a linear partial differential operator L_{BL} acting on the vector space of homogeneous polynomials in three indeterminates. The method is applicable to other Two Urns models and is relatively easy to use compared to previous methods based on orthogonal polynomials or group representations.

Key word. Urn modelsDiffusionEigenvalues Generating Functions

1. Introduction. The Bernoulli-Laplace (BL) model arise from diffusion theory and is related to the shuffling of cards [8]. Symmetries of the permutation group S_N appear naturally in this model and other random walks on groups. Previous solutions of this model have appeared in Diaconis and Shashahani [7] and in the works of Karlin and MacGregor [10]. Group representations are used explicitly in the first; the derivation of a non-standard inner product or equivalently a measure for orthogonal polynomials which are related to the eigensolutions of the BL model appears in the second.

In this paper, we give a third way for deriving exact solutions of all the eigenvectors of the BL model, through a nonlinear transform that triangularizes and then diagonalizes the transition matrix T_{BL}. In brief, our method associates a specific linear partial differential operator (LPDO) L_{BL} that acts on the vector space of homogeneous polynomials, G, to the matrix T_{BL}. The L_{BL} inherits the symmetries of the BL model; it encodes the tri-diagonal singly-stochastic (column sums are equal to 1) and anti-symmetric structure of T_{BL}. The components of the (right) eigenvectors (in view of the equal column sums of T_{BL}) of T_{BL} is encoded in the coefficients of the homogeneous polynomial G. A classical theory for the symmetries of such LPDOs have been formulated in terms of the Lie algebra of symmetry operators K that commutes with L_{BL} (cf. [13]).

It turns out and we exploit in our method, that the symmetries of L_{BL} appear iu the form of suitable linear and nonlinear transformations P on the independent variables x, y, etc. or indeterminates of G. The ease of use of this method resides in the transparent or explicit way to find these transformations P. Our algorithm is completed by associating the transformed LPDO, L'_{BL},back to what turns out to be a triangular matrix T'_{BL}; in other words, the transformation P for L_{BL} encodes a similarity transformation that triangularizes T_{BL}, i.e., $PT_{BL}P^{-1} = T'_{BL}$, which is then solved directly for its eigenvalues and right eigenvectors.

Here, we give a summary of the Urn models to which the BL model is related as an extension. The Ehrenfest model and the Polya Urn models are two of the early solvable models in the literature [9]. They appear as two of the exactly solved cases in Friedman's formulation of Urn models where precisely one urn and balls of two colors are drawn and replaced with additions [9]. A dual formulation of Friedman's Urn models was introduced in a series of recent papers [16, 14]: instead of balls of two colors and one urn, the dual formulation uses two urns and one-colored balls. The latter is more convenient for modelling of certain network science models [3], such as the Voter model where two balls are drawn and returned to the two urns with prescribed probabilities that depend on the order in which they are drawn. This is because many of the network science models are irreversible Markov chain models [4] which have absorbing states. Their transition matrices T, unlike T_{BL} for the BL model, are not symmetric, in an essential sense, that is, there are no non-standard inner products for R^N in which these matrices T have a symmetric form.

Using a new method based on diagonalization of transition matrices [15], we solved exactly the eigenvectors of several well-known models, including the Ehrenfest model, the Voter model [18, 12, 11, 5, 17, 2], the Moran model for genetic drift, and the Naming game models [19], [20]. Most of these models are irreversible Markov models with absorbing states, and have essentially non-symmetric transition matrices in the sense just mentioned. The BL model however, is based on two urns and balls of two colors. Thus, it is not strictly in the class of Two Urns models to which we recently applied our method. In modifying this method so that it applies to the BL model, we will have shown that the new method is not only easy to use but also flexible in extension to new problems.

One of the main points here is the technical simplicity of uncovering the symmetries of the above LPDOs within our method, through the explicit appearance of the expressions $u = f(x, y)$ in the coefficients of L. We give here the LPDO L_V for the Voter model [16, 14] to indicate what we mean: first the propagation equation for the transition matrix T_V is given by

$$a_j^{(m+1)} = p_{j-1} a_{j-1}^{(m)} + (1 - 2p_j) a_j^{(m)} + p_{j+1} a_{j+1}^{(m)} \tag{1.1}$$

$$p_j = \frac{j(N-j)}{N(N-1)}; \tag{1.2}$$

the eigen-problem for the associated LPDO $L_V = (x - y)^2 G_{xy}$ acting on the homogeneous polynomial $G(x, y) = \sum_j c_j x^j y^{N-j}$ (which encodes the components c_j of the right-eigenvector of T_V) is given

$$(x - y)^2 G_{xy} = N(N-1)(\lambda - 1)G, \tag{1.3}$$

which clearly suggests the transformation $u = x - y$, $v = y$. Indeed this triangularized and diagonalized the Voter model and led to its complete solution.

Contrast this ease of use with the fact that triangularization and diagonalization of a given transition matrix of size N has computational complexity $O(N^3)$. In other words, exact integration of the Two Urns models via diagonalization of transition matrices are nontrivial problems, that are difficult to solve but once known, the solutions are easy to verify. [16, 14] provides a simple method to find such explicit diagonalization and hence all eigenvectors for a class of transition matrices from the Two Urns models, even when their transition matrices are essentially non-symmetric. Note that the eigenproblem and diagonalization of symmetric matrices have a lower computational complexity.

We aim here to highlight this method's ease of use, relative to the group representation method and the method of orthogonal polynomials. Moreover, the BL model differs significantly from the original Two Urns subclass of models for which our method was initially formulated. Thus, we also aim to show that, with the specific introduction of a nonlinear change of independent variables, this method can be applied to more complex models than the original class of models. Since the BL transition matrix T_{BL} is from a reversible Markov chain with a stationary distribution [4], it is non-symmetric singly-stochastic only in a trivial sense. In other words, there exists (a difficult to find) non-standard inner product for R^N, in which T_{BL} is symmetric and hence doubly-stochastic. Karlin and McGregor [10] using their powerful integral representation method of finding an explicit way to symmetrize T by introducing a non-standard inner product (or orthogonal measure) into the problem, have related the right eigenvectors of the transition matrix T_{BL} of the BL model to the orthogonal polynomials called the dual Hahn polynomials. A third aim of this paper is therefore to re-derive from the diagonalization of T_{BL}, this non-standard inner product in which the dual Hahn polynomials are an orthogonal polynomial system. Note that this non-standard inner product, once found, yields a symmetric version of T_{BL}which is an example of a Jacobi operator that arise in the the classical moments problem [1], and is related to orthogonal polynomials via the Riemann-Hilbert method [6].

The paper is organized as follows: section 3 concerns the calculation of the right eigenvectors and eigenvalues of T in closed form by method introduced in [16, 15, 14]; in view of the fact that these right eigenvectors are not orthogonal in the usual Euclidean inner-product, section 4 concerns the transformations needed to calculate the orthonormal system of left eigenvectors of T_{BL},and also the derivation of the non-standard inner product in which the right eigenvectors are now orthogonal; section 5 concerns the elementary proofs, based only on the eigenvectors and eigenvalues of T, for the tight upper and lower bounds for times to stationarity in the BL model [7] ; section 6 concerns a numerically exact evaluation of the expression for the TV norm in these bounds using the eigenvectors and eigenvalues of T_{BL} directly, hence slightly sharper estimates for the mixing times of the BL model.

Beyond the balanced special case of the BL problem treated in detail in this paper, the same generating function method can be used to prove similar tight bounds for mixing times in the other cases.

2. Transition matrix of the BL model. Let the transition matrix T_{BL} be defined so that $(T_{BL}\ )_{ij} = Pr\{n_w(m+1) = i \mid n_w(m) = j\}$, so that the sum of each column is 1. In the general BL model for balls of two colors and two urns, N_1, N_2, N_w, N_b are fixed parameters satisfying the constraints

$$N_w + N_b = N = N_1 + N_2 \tag{2.1}$$

where N equals total number of balls in the model. For $i = 0, ..., N_w \le N_1$, (where by abuse of notation i stands for both the row label of transposed matrix T^t_{BL} and the number of white balls in urn 1, n_w), the transition probabilities are explicitly given by

$$p_i = \Pr\{n_w(t+1) = i+1 |\ n_w(t) = i\} = \frac{(N_1 - i)(N_w - i)}{N_1 N_2} \tag{2.2}$$

$$q_i = \Pr\{n_w(t+1) = i-1 |\ n_w(t) = i\} = \frac{i\,(N_b - (N_1 - i))}{N_1 N_2} \tag{2.3}$$

$$r_i = \Pr\{n_w(t+1) = i |\ n_w(t) = i\} = 1 - q_i - p_i. \tag{2.4}$$

.

3. Diagonalization - Right eigenvectors of the general BL model. In [16, 14], we developed an explicit method for exactly integrating or solving a 5-parameters subclass of a class of Two Urns models which is parametrized by six real parameters. Our method is based on a relationship between certain banded stochastic matrices T (such as tridiagonal and pentadiagonal non-symmetric transition matrices of markov chain models) and the LPDOs acting on the vector space of homogeneous polynomials, $G(x, y)$ of finite order in two indeterminates. The symmetries of the LPDO, L, associated with a given non-symmetric singly stochastic matrix from this solvable subclass of the Two Urns models, are identified and used explicitly to transform from the original indeterminates (independent variables x, y say) to suitable new variables (such as $u = f(x, y), v = g(x, y)$). In the new variables u, v, the transformed LPDO, L', acts on the (again homogeneous of same order as $G(x, y)$) polynomial $H(u, v)$. We have shown in [16, 14] that the transformed eigen-problem

$$L'[H(u, v)] = \lambda(N)H(u, v) \tag{3.1}$$

for a well-defined subclass of such Two Urns problems is equivalent (via the inverse of the original relationship between banded matrix and LPDO) to the eigen-problem for a triangular matrix, which can then be solved explicitly for both right and left eigenvectors. In other words, at the end of this brief summary, the symmetries of LPDO L inherited from the original banded stochastic matrix T,generate an explicit similarity transformation, P, such that

$$STS^{-1} = D \tag{3.2}$$

where D is diagonal, and S contains the eigenvectors of T.

This method can be formalized as an Algorithm as follows: Given the input of a singly stochastic transition matrix T of size $N + 1$,

(I) Choose a suitable homogeneous polynomial of finite degree N, G that has the components c_i of a right eigenvector of T as coefficients of the monomials $x^i y^j z^k$; part of this choice is the number of indeterminates in G. For example, the Voter model of size N (number of balls) with a transition matrix $T_V(N)$ which is a $N + 1$ by $N + 1$ real matrix, requires a homogeneous polynomial G_V of degree N in the indeterminates x, y because there are two urns.

(II) Associate the recursion inplicit in given Markov matrix T to a LPDO, L which acts on the homogeneous polynomial G; the basic elements of this association scheme are the standard linear differential operators for increasing, decreasing and not changing the numbers of balls in each urn (which correspond in the example below to the probabilities p, q, r prescibed by the transition matrix), and a set of multiplication type linear operators that correspond to shifts.

(III) A transformation to new independent variables,(for instance, $u = f(x, y, z)$, $v = g(x, y, z)$, $w = h(x, y, z)$) is chosen to satisfy two conditions:

(A) the transformed polynomial

$$H(u, v, w) = H(f(x, y, z), g(x, y, z), h(x, y, z)) = G(x, y, z)$$

is a homogeneous polynomial of the same finite degree as G;

(B) $u = f(x,y,z)$, $v = g(x,y,z)$, $w = h(x,y,z)$ is a transformation based on the symmetries of L (cf. [13]), that is, the combinations $f(x,y,z)$, $g(x,y,z)$, $h(x,y,z)$ of the original variables x, y, z appear naturally in the coefficients of the LPDO, L.

These conditions (A) and (B) are clearly not sufficient to ensure the transformed LPDO eigenproblem

$$L'[H(u,v,w)] = \lambda(N)H(u,v,w) \tag{3.3}$$

is associated with a similar triangular matrix T' which explicitly yields all its eigenvectors b_i. That they are sufficient has to be proved either in each problem to which we apply the Algorithm, or for a class of models as in the case of the Two Urns models.

(IV) Using the transformation in step (III), derive the corresponding transformed LPDO, L' that acts on the transformed polynomial $H(u,v,w)$.

(V) Without explicitly calculating the transformed matrix T' which is associated with the transformed LPDO, L' in step (IV), check that the transformed eigenproblem for L' is indeed a recursion system for the transformed eigenvectors b_i that can be solved explicitly, i.e., it is equivalent to a triangular linear system of equations. Solve for the eigenvalues and then the eigenvector components b_i, and if required transform back to the original components c_i. These are the main outputs of the Algorithm.

(VI) Use the eigenvectors in step (V) to diagonalize the original matrix if necessary. This is the end of the ALgorithm.

Now we apply the Algorithm to the BL model. Given the transition matrix of the BL model (cf. section 2), it will be obvious that three independent variables (instead of the two before) should be used to formulate the BL problem. In step (I) of the Algorithm, we adopt the anzatz that the $LPDO$, L_{BL}, associated with the above N by N matrix T_{BL}, now acts on a homogeneous polynomial $G(x,y,z)$ in three indeterminates, x, y, z. We encode the entries $c_k(i)$, $i = 0, ..., N_w$ of the $k-th$ eigenvector of the transition matrix for the BL model as follows:

$$G^{(k)}(x,y,z) = \sum_{i=0}^{N_w} c_k(i) x^i y^{N_1 - i} z^{N_w - i}. \tag{3.4}$$

where i = number of white balls in urn 1 (also denoted n_w). The choice of three independent variables to encode the components of an eigenvector of T_{BL} in the homogeneous polynomial $G(x,y,x)$ is now made obvious by this explicit expression for G.

In step (II) of the Algorithm, we derive from the original eigen-problem for transition matrix T_{BL}, an LPDO, L_{BL}, that acts on $G^{(k)}$. Towards that aim, we note, in particular, the entries for p_i and q_i in T_{BL} correspond respectively to the following linear differential operators with coefficients that are monomials in x, y,and z,

$$L_p = \frac{yz G^{(k)}_{yz}}{N_1 N_2} \tag{3.5}$$

$$L_q = \frac{N_b x}{N_1 N_2} G^{(k)}_x - \frac{xy}{N_1 N_2} G^{(k)}_{xy}, \tag{3.6}$$

where $G_{yz}^{(k)} = \frac{\partial^2}{\partial y \partial z} G^{(k)}$ for example. In addition, it is part of the association scheme that multiplication in the LPDO (cf. [13]) by the coefficient $\frac{x}{yz}$ (resp. $\frac{yz}{x}$) represents down (resp. up) shifts in the index i within the discrete recursion equations of the original eigen-problem for matrix $T_{BL.}$ The L_{BL} associated with the eigen-problem of the tridiagonal Markov matrix $T_{BL}$is given by:

$$L_{BL}[G^{(k)}] = N_1 N_2 (\lambda_k - 1) G^{(k)} \tag{3.7}$$

$$L_{BL}[G^{(k)}] \equiv (x - yz) G_{yz}^{(k)} + y (x - yz) G_{xy}^{(k)} - N_b (x - yz) G_x^{(k)}. \tag{3.8}$$

In step(III) of the Algorithm, we note that the symmetries of L_{BL} with respect to transformations of its independent variables, is expressed in the factor $(x - yz)$ in its coefficients. This suggests the transformation to the new independent variables

$$u = x - yz, y = y, z = z. \tag{3.9}$$

Since the transformed homogeneous polynomial is now given by

$$H^{(k)}(u, y, z) = G^{(k)}(x(u, yz), y, z) = \sum_i b_i^k u^i y^{N_1 - i} z^{N_w - i} \tag{3.10}$$

in terms of the (new) components b_i^k of the $k-th$ right eigenvector, this transformation clearly satisfies both necessary conditions (A) and (B) in step (III) of the Algorithm.

To prove that it is sufficient for our purpose of obtaining the eigenvalues and eigenvectors exactly and for diagonalization, we proceed by direct calculations.

In step (IV) using the following obvious identities for the transformation of partial derivatives

$$\partial_x = \partial_u \tag{3.11}$$

$$\partial_y = \partial_y - z\partial_u \tag{3.12}$$

$$\partial_z = \partial_z - y\partial_u \tag{3.13}$$

$$\partial_{xy} = \partial_{yu} - z\partial_u^2 \tag{3.14}$$

$$\partial_{yz} = \partial_{yz} - y\partial_{yu} - z\partial_{uz} + yz\partial_u^2 - \partial_u \tag{3.15}$$

the transformed LPDO, L'_{BL} in $H^{(k)}$, $k = 0, ..., N_w$ is

$$N_1 N_2 (\lambda_k - 1) H^{(k)} = L'_{BL}[H^{(k)}], \tag{3.16}$$

$$L'_{BL}[H^{(k)}] = -N_b u \partial_u H^{(k)} + yu \left(\partial_{yu} - z\partial_u^2\right) H^{(k)} \tag{3.17}$$

$$+u \left(\partial_{yz} - y\partial_{yu} - z\partial_{uz} + yz\partial_u^2 - \partial_u\right) H^{(k)} \tag{3.18}$$

$$= u \left(\partial_{yz} - z\partial_{uz}\right) H^{(k)} - (N_b + 1) u \partial_u H^{(k)} \tag{3.19}$$

In step (V), by reversing the derivation of the original L_{BL} through the association scheme [13], this L'_{BL} in H is shown to be equivalent to the following triangular system for the (right) eigen-problem of the transformed matrix T'_{BL}:

$$N_1 N_2 (\lambda_k - 1) b_i^k \tag{3.20}$$

$$= (N_1 - i + 1) (N_w - i + 1) b_{i-1}^k - i (N_w - i) b_i^k - (N_b + 1) i b_i^k. \tag{3.21}$$

We have therefore verified the sufficiency of the transformation where $u = x - yz$ for triangularizing (and later diagonalizing) T_{BL}. This triangular system implies the recursion

$$b_i^k = \frac{(N_1 - i + 1) (N_w - i + 1) b_{i-1}^k}{N_1 N_2 (\lambda_k - 1) + i (N_w - i) + (N_b + 1) i} \tag{3.22}$$

which can be solved directly.

For nontrivial eigensolutions for $k = 0, ..., N_w$, the denominator in b_i^k must vanish, yielding the following exact expressions for the eigenvalues,

$$\lambda_k = 1 - \frac{k\,(1 - k + N_w + N_b)}{N_1 N_2} \tag{3.23}$$

$$= 1 - \frac{k\,(1 - k + N)}{N_1 N_2} \tag{3.24}$$

$$= 1 - \frac{k\,(N - k + 1)}{N_1 N_2} \tag{3.25}$$

In the case $N_1 = N_w$,

$$\lambda_0 = 1 \tag{3.26}$$

$$\lambda_1 = 1 - \frac{N}{N_1 N_2}. \tag{3.27}$$

The eigenvectors (in the transformed variables of H) are given explicitly by:

$$b_i^k = \prod_{j=k+1}^{i} \frac{(N_1 - j + 1)\,(N_w - j + 1)}{N_1 N_2 (\lambda_k - 1) + j\,(N_w + N_b + 1 - j)} \tag{3.28}$$

$$= \prod_{j=k+1}^{i} \frac{(j - N_1 - 1)\,(j - N_w - 1)}{-k\,(N + 1 - k) + j\,(N + 1 - j)} \tag{3.29}$$

$$= \prod_{j=k+1}^{i} -\frac{(j - N_1 - 1)\,(j - N_w - 1)}{(j - k)\,(j + k - N - 1)} \tag{3.30}$$

$$= (-1)^{i-k} \frac{(k - N_1)_{i-k} (k - N_w)_{i-k}}{(i - k)!(2k - N)_{i-k}}. \tag{3.31}$$

Using these coefficients in the definition for H gives

$$H^{(k)} = \sum_{i=k}^{N'} (-1)^{i-k} \frac{(k - N_1)_{i-k} (k - N_w)_{i-k}}{(i - k)!(2k - N)_{i-k}} u^i y^{N_1 - i} z^{N_w - i} \tag{3.32}$$

We summarize the consequences of the above steps of the Algorithm on the BL model in the following theorem:

Theorem 3.1. *In the above Algorithm for the BL model, for any size N of the model, the LPDO, L'_{BL}, after the transformation (3.9) on the independent variables, is equivalent to a triangular linear system (3.21) which has (right) eigenvectors given by (3.31) and eigenvalues (3.25). The (right) eigenvectors of the original BL matrix T_{BL} are in turn given by (3.38).*

3.1. Hypergeometric functions and dual Hahn polynomials.

The next to final step left in this part of the paper is step (VI) in the Algorithm, to invert the above similarity transformation to obtain explicitly the closed-form expressions for the original components of the right-eigenvectors $c_k(i)$ of $T_{BL.}$ For this purpose, let $h^{(k)}(u) = H^{(k)}(u, 1, 1)$. Then,

$$\begin{aligned} g^{(k)}(x) = G^{(k)}(x, 1, 1) = H^{(k)}(x - 1, 1, 1) = h^{(k)}(x - 1) \\ = (x - 1)^k {}_2F_1(k - N_1, k - N_w; 2k - N; 1 - x). \end{aligned}$$

Using the hypergeometric identity [],

$$ {}_2F_1(a,b;c;1-z) \propto {}_2F_1(a,b;a+b-c+1;z) \tag{3.33} $$

and the fact that any multiple of an eigenvector remains an eigenvector, we take the polynomial for the right eigenvector components to be

$$ g^{(k)}(x) = (x-1)^k {}_2F_1(k-N_1, k-N_w; N_2-N_w+1; x) \tag{3.34} $$

whose coefficients are the original components $c_k(i)$ of the $k-th$ right eigenvector corresponding to λ_k prior to the transformation above. These expressions are equivalent to the dual Hahn polynomials [10].

For the hypergeometric representation of the eigenvectors to be well defined, we require $N_2 \geq N_w$. There is no loss in generality with this assumption, because we can relabel $N_1 \leftrightarrow N_2$ and $N_w \leftrightarrow N_b$ so that the assumption holds. From the solution for $g^{(k)}$, we expand in x^l,

$$ g^{(k)}(x) = \sum_n \binom{k}{n} (-1)^{k-n} x^n \sum_i \frac{(k-N_1)_i (k-N_w)_i}{(N_2-N_w+1)_i i!} x^i \tag{3.35} $$

$$ = \sum_i \sum_n \binom{k}{n} (-1)^{k-n} \frac{(k-N_1)_i (k-N_w)_i}{(N_2-N_w+1)_i i!} x^{i+n} \tag{3.36} $$

$$ = (-1)^k \sum_i \left[\sum_n \binom{k}{n} (-1)^n \frac{(k-N_1)_{i-n} (k-N_w)_{i-n}}{(N_2-N_w+1)_{i-n} (i-n)!} \right] x^i \tag{3.37} $$

to find the explicit form for the components of the $k-th$ right eigenvectors:.

$$ c_i^k = \sum_n \binom{k}{n} (-1)^n \frac{(k-N_1)_{i-n} (k-N_w)_{i-n}}{(N_2-N_w+1)_{i-n} (i-n)!} \tag{3.38} $$

Notice that the solution is the kth order backwards difference of the components of the hypergeometric coefficients of g.

The above treatment of the eigen-problem by transforming via symmetries, the independent variables of the associated LPDO, L_{BL}, is equivalent to a similarity transformation of the transition matrix T_{BL}. Let $\mathbf{w} = \mathbf{P}\mathbf{v}$ for some transformation matrix $\mathbf{P}$. Then, the eigen-problem for $\mathbf{w}$ is given by $\mathbf{P}\mathbf{T}_{BL}\mathbf{P}^{-1}\mathbf{w} = \lambda\mathbf{w}$. The above calculations is equivalent to the matrix $\mathbf{P}$ such that the new matrix $\mathbf{T}'_{BL} = \mathbf{P}\mathbf{T}\mathbf{P}^{-1}$ is lower triangular. The last step in this section is to diagonalize $T_{BL.}$. We do this by diagonalizing the matrix triangular matrix $\mathbf{T}'_{BL} = \mathbf{W}\mathbf{\Lambda}\mathbf{W}^{-1}$. Here, $\mathbf{\Lambda} = diag(\lambda_0, \ldots, \lambda_N)$ and $\mathbf{W}$ are the eigenvectors of $\mathbf{T}'_{BL}$. The components of these eigenvectors are b_i corresponding to eigenvalue λ_k. Diagonalization of $\mathbf{T}'_{BL}$ allows us to explicitly diagonalize the original transition matrix as

$$ \mathbf{T}_{BL} = \mathbf{P}^{-1}\mathbf{W}\mathbf{\Lambda}\mathbf{W}^{-1}\mathbf{P}. \tag{3.39} $$

Note that the matrix of eigenvectors is given by $\mathbf{P}^{-1}\mathbf{W}$.

4. Symmetrizing transform, orthogonal measure and dual Hahn polynomials. For transition matrix T_{BL}, let Z be given by (where we drop the subscript BL herein, i.e., $T = T_{BL}$)

$$ Z_{ij} = \sqrt{\pi_j} T_{ij} \frac{1}{\sqrt{\pi_i}}. \tag{4.1} $$

Recall the detailed balance of T and its stationary distribution, given by $T_{ij}\pi_j = T_{ji}\pi_i$ [4] follows from the reversibility and ergodicity of the BL model. Note that Z is the symmetric version of the transition matrix:

$$Z_{ij} = \frac{1}{\sqrt{\pi_j}}\pi_j T_{ij}\frac{1}{\sqrt{\pi_i}} \tag{4.2}$$

$$= \frac{1}{\sqrt{\pi_j}}T_{ji}\sqrt{\pi_i} \tag{4.3}$$

$$= Z_{ji}. \tag{4.4}$$

Therefore, Z has an orthonormal set of left eigenvectors, w_k^T. Let W be a matrix whose columns are w_k. The spectral decomposition of Z by left eigenvectors is given by

$$Z = W\Lambda W^T. \tag{4.5}$$

By the definition of Z_{ij}, the transformation from T to Z can be expressed as

$$D^{-1}TD = Z, \tag{4.6}$$

where D is a diagonal matrix whose diagonal entries are $\sqrt{\pi_i}$. So, arbitrary powers of T is given by

$$T^m = DW\Lambda^m(D^{-1}W)^T. \tag{4.7}$$

Defining a new transformation by S,

$$T^m = S\Lambda^m S^{-1}. \tag{4.8}$$

Since W has a specific normalization, we can equate S with DW after applying the appropriate normalization for S. That is, for diagonal matrix Δ, we take

$$S\Delta = DW. \tag{4.9}$$

We can choose any normalization for the right eigenvectors given in S, and Δ will properly renormalize them. Here, we solve for Δ and W by appealing to the orthogonality of W:

$$W^TW = \Delta S^T D^{-2} S\Delta = I. \tag{4.10}$$

Therefore

$$S^T D^{-2} S = \Delta^{-2}. \tag{4.11}$$

Computing the matrix multiplication on the left side yields the diagonal entries of Δ denoted by Δ_k given by

$$\Delta_k^{-2} = \sum_{i=0}^{N}\frac{1}{\pi_i}c_k(i)^2 \tag{4.12}$$

in terms of the right eigenvectors of T_{BL}. Now that we have Δ, we have the representations for both the left-eigenvectors $w_k(i)$ and right-eigenvectors $v_k(i)$ of Z given by

$$w_k(i) = \frac{\Delta_k}{\sqrt{\pi_i}} c_k(i) \tag{4.13}$$

$$v_k(i) = \frac{1}{\sqrt{\pi_i}} w_k(i) = \frac{\Delta_k}{\pi_i} c_k(i) \tag{4.14}$$

in terms of the right-eigenvectors $c_k(i)$ of the original T_{BL} that was obtained by our method in section 2.

From Eq. (4.11), we also have an explicit formula for S^{-1} given by

$$S^{-1} = \Delta^2 S^T D^{-2}. \tag{4.15}$$

So, by Eq. (**??**), we have

$$T^m = S\Lambda^m \Delta^2 S^T D^{-2}. \tag{4.16}$$

Computing the matrix multiplication gives the following spectral decomposition

$$T_{ij}^{(m)} = \frac{1}{\pi_j} \sum_{k=0}^{N} \Delta_k^2 \lambda_k^m c_k(i) c_k(j). \tag{4.17}$$

as the explicit representation of $Pr\{n(m) = i \mid n(0) = j\}$ in the BL model.

Since $T^0 = I$, take $m = 0$ in Eq. (4.17) to find the stationary distribution of the BL model

$$\pi_j \delta_{ij} = \sum_{k=0}^{N} \Delta_k^2 c_k(i) c_k(j). \tag{4.18}$$

Note that this is the orthogonality relation for the right-eigenvectors of T_{BL} with orthogonal measure Δ_k^2 given in Eq. (**??**). We have derived the orthogonal measure Δ_k^2 in which the dual Hahn are an orthogonal polynomial system [10].

We summarize these results on the derivation of a non-standard inner product or orthogonal measure in which the original transition matrix of the BL model becomes a symmetric real matrix and the (right) eigenvectors are the system of orthogonal dual Hahn polynomials:

THEOREM 4.1. *The orthogonal measure in (4.12) symmetrizes T_{BL}, is related to the (left) and (right) eigenvectors of T_{BL} by (4.13) and (4.14), and yields the spectral decomposition (4.17).*

5. Bounds of mixing times - elementary proofs. We will discuss first the

case $N_1 = N_2 = N/2$, for our method gave the eigenvalues of the BL model to be $\lambda_k = 1 - \frac{4k(N-k+1)}{N^2}$, $\lambda_1 = 1 - \frac{4}{N}$. A heuristic estimate of the number of switches q needed to mix the colors is thus,

$$(1 - \frac{4}{N})^q \simeq e^{-\frac{4q}{N}} = \frac{1}{N}, \tag{5.1}$$

and therefore $q = \frac{1}{4} N \log N$ gives an idea of how many switches or time steps are needed until the variation distance between ρ_q and π is of order $O(1/N)$. The lower bound is obtained along the lines of Diaconis et al, that is, by an application of the Chebyshev's inequality. However, all estimates of the relevant mean values and variance needed to apply the Chebyshev's inequality are constructed explicitly from the properties of the eigenvalues and eigenvectors of the BL model. We will state the theorems below but since the proofs are similar to those in [7], we have provided the details in an appendix.

THEOREM 5.1. *For* $m = \frac{1}{4} N \log N + (\frac{c}{2} - \frac{\log 2}{8})N$, *for* $c > 0$, *there is a universal constant* $A > 0$ *such that* $E_\pi \left[\|\rho_m(j;.) - \pi\|_V \right] \le A e^{-2c}$.

THEOREM 5.2. *If* $m = \frac{1}{8} N \ln N - \frac{cN}{2}$, *then* $2||\rho_m - \pi||_V \ge 1 - e^{4c}$

6. Exact calculations of mixing times. Given that we can calculate $\rho_m(i,j)$ exactly by Eq. (**??**), and the stationary distribution is given by

$$\pi_i = \frac{\binom{N_w}{i}\binom{N-N_w}{N_1-i}}{\binom{N}{N_1}}, \tag{6.1}$$

the total variational distance can be exactly computed for all time steps by

$$\|\rho - \pi\|_V = \frac{1}{2} \sum_{i=0}^{N/2} \left| \pi_i \sum_{k=1}^{N/2} \lambda_k^m v_k(i) v_k(0) \right| \tag{6.2}$$

This solution is shown in Figure 1, with the upper bound given in Figure 2.

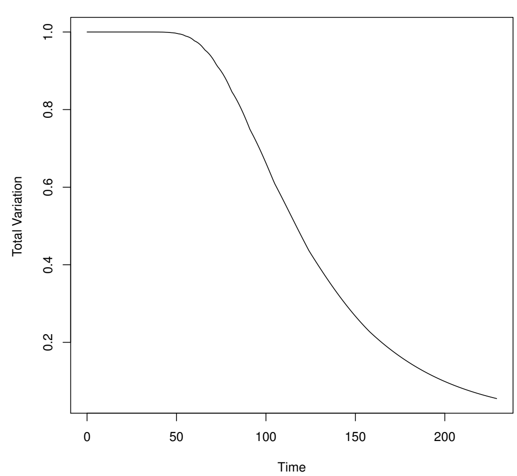

FIG. 1. *Plot of the exact solution for the total variational distance for* $N_1 = N_2 = N_w = N_b = 100$.

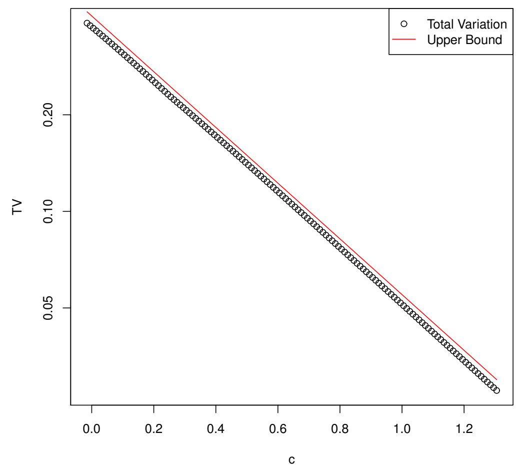

FIG. 2. *Plot of the exact solution for the total variational distance, and the upper bound given in Theorem 5.1*

Appendix A. Proofs of Theorems 5.1 and 5.2. The eigenvalues and eigenvectors of the BL problem can be used to construct an upper bound on the variation distance between *the* probability distribution after m. As we proved in section 3, the right-eigenvectors of the original matrix T_{BL} are not orthogonal in the standard inner product of R^N but a different inner product weighted by Δ_k^2 can be used to derive orthogonality of a related system of right-eigenvectors v_j of $T = T_{BL}$.

A.1. Upper bound. Let $j = 0$ to define

$$\rho_m(i;0) = \pi_i + \pi_i \sum_{k=0}^{N/2} \lambda_k^m v_k(i) v_k(0)$$

where

$$\Pr\{n_w(m) = i \mid n_w(0) = j\} = T_{ij}^m = \rho_m(j;i) \tag{A.1}$$

$$= \sum_{k=0}^{N/2} \pi_i \lambda_k^m v_k(i) v_k(j) \tag{A.2}$$

$$= \pi_i + \pi_i \sum_{k=1}^{N/2} \lambda_k^m v_k(i) v_k(j). \tag{A.3}$$

Then, $\rho_0(i;j) = \delta_{ij}$, and hence

$$\sum_{k=1}^{N/2} v_k^2(i) = \frac{1}{\pi_i} - 1 < \frac{1}{\pi_i}, \tag{A.4}$$

implies that

$$\text{(A.5)}\quad \|\rho_m(i;0)-\pi_i\|_V = \frac{1}{2}\sum_{i=0}^{N/2}\left|\pi_i\sum_{k=1}^{N/2}\lambda_k^m v_k(i)v_k(0)\right|$$

$$\text{(A.6)}\quad \leq \frac{1}{2}\lambda_1^m\sum_{i=0}^{N/2}\pi_i\left|\sum_{k=1}^{N/2}v_k(i)v_k(0)\right|$$

$$\text{(A.7)}\quad \leq \frac{1}{2}(N)^{-1/2}e^{-2c}\sum_{i=0}^{N/2}\pi_i\left(\sum_{k=1}^{N/2}v_k^2(i)\right)^{1/2}\left(\sum_{k=1}^{N/2}v_k^2(0)\right)^{1/2}$$

$$\text{(A.8)}\quad \leq \frac{1}{2}(N\ \pi_0)^{-1/2}e^{-2c}\sum_{i=0}^{N/2}(\pi_i)^{1/2}$$

Using

$$\text{(A.9)}\quad \sum_{i=0}^{N/2}(\pi_i)^{1/2} = O(N^{1/4})$$

and averaging over initial data $n_w(0)=j$, $j=0,...,N/2$,we have

$$\text{(A.10)}\quad E_\pi\left[\|\rho_m(j;i)-\pi_i\|_V\right] \leq \sum_{j=0}^{N/2}\pi_j\,\|\rho_m(j;.)-\pi\|_V$$

$$\text{(A.11)}\quad \leq \frac{1}{2}N^{-1/2}e^{-2c}\sum_{j=0}^{N/2}(\ \pi_j)^{1/2}\sum_{i=0}^{N/2}(\pi_i)^{1/2}$$

$$\text{(A.12)}\quad \leq Ae^{-2c}$$

for some A independent of N. This proves theorem 5.1.

A.2. Lower bound. In terms of the right-eigenvectors v_k, $k=0,...,N/2$, (with col sum =1)

$$\text{(A.13)}\quad \Pr\{n_w(m)=i \mid n_w(0)=j\} = T_{ij}^m = \rho_m(i) \text{ if we take } j=0$$

$$\text{(A.14)}\quad = \sum_{k=0}^{N/2}\pi_i\lambda_k^m v_k(i)v_k(j)$$

$$\text{(A.15)}\quad = \pi_i + \pi_i\sum_{k=0}^{N/2}\lambda_k^m v_k(i)v_k(j)$$

and

$$E_{\rho_m}[v_1(i)] = \sum_{i=0}^{N/2} v_1(i)\rho_m(i) = \sum_{i=0}^{N/2} \pi_i v_1(i) \sum_{k=0}^{N/2} \lambda_k^m v_k(i) v_k(0) \tag{A.16}$$

$$= \sum_{i=0}^{N/2}\sum_{k=0}^{N/2} \lambda_k^m \pi_i v_1(i) v_k(i) v_k(0) \tag{A.17}$$

$$= \sum_{k=0}^{N/2} v_k(0)\lambda_k^m \sum_{i=0}^{N/2} \pi_i v_1(i) v_k(i) \tag{A.18}$$

$$= v_1(0)\lambda_1^m \sum_{i=0}^{N/2} \pi_i v_1(i) v_1(i) \tag{A.19}$$

$$= v_1(0)\lambda_1^m = v_1(0)\left(1-\frac{4}{N}\right)^m \tag{A.20}$$

$$E_\pi[v_1(i)] = 0; var_\pi\{v_1(i)\} = 1. \tag{A.21}$$

Next for $m = \frac{1}{8}N\log N - c\frac{N}{2}$, we get $E[v_1] = \frac{v_1(0)}{\sqrt{N}}e^{2c}$. A similar calculation gives

$$E_{\rho_m}[v_2(i)] = \sum_{i=0}^{N/2} v_2(i)\rho_m(i) = \sum_{i=0}^{N/2} \pi_i v_2(i) \sum_{k=0}^{N/2} \lambda_k^m v_k(i) v_k(0) \tag{A.22}$$

$$= \sum_{i=0}^{N/2} \pi_i v_2(i) \sum_{k=0}^{N/2} \lambda_k^m v_k(i) v_k(0) = \sum_{i=0}^{N/2}\sum_{k=0}^{N/2} \lambda_k^m \pi_i v_2(i) v_k(i)\ v_k(0) \tag{A.23}$$

$$= \sum_{k=0}^{N/2} v_k(0)\lambda_k^m \sum_{i=0}^{N/2} \pi_i v_2(i) v_k(i) \ = \ v_2(0)\lambda_2^m \sum_{i=0}^{N/2} \pi_i v_2(i) v_2(i) \tag{A.24}$$

$$= v_2(0)\lambda_2^m \sim v_2(0)\left(1-\frac{8}{N}\right)^m, \tag{A.25}$$

$$E_\pi[v_1(i)] = 0;\ var_\pi\{v_1(i)\} = 1 \tag{A.26}$$

Next we deduce $v_1(0)$ and $v_2(0)$ from $v_1^2 = Av_2 + B$, $v_1(i) = C(N/4 - i)$, and the orthogonality of v_i:

$$1 = \sum_{i=0}^{N/2} v_1(i)^2\pi_i = \sum_{i=0}^{N/2} C^2(i - N/4)^2\pi_i \sim C^2\frac{N}{16} \tag{A.27}$$

Therefore, taking $C = \frac{4}{\sqrt{N}}$, we fine $v_1(0) \sim \sqrt{N}$. Furthermore, $Av_2(0) = v_1(0) - b$. Now, we have

$$Var_{\rho_m}\{v_1\} = E_{\rho_m}[Av_2 + B] - N\lambda_1^{2m} \tag{A.28}$$

$$= (N-B)\lambda_2^m + B - N\lambda_1^{2m} \sim B(1-\lambda_2^m) \tag{A.29}$$

So with the same normalization as above for v_1, we deduce $Var\{v_1'\}$ is uniformly bounded by constant $2b$, since $B = b + O(\log N/N)$. Now, by Chebyschev's inequality,

$$Pr_\pi\{|v_1| \le k\} \ge 1 - \frac{1}{k^2} \tag{A.30}$$

and

$$Pr_{\rho_m}\{|v_1| \leq k\} \leq Pr_{\rho_m}\{v_1 \leq k\} \tag{A.31}$$
$$= Pr_{\rho_m}\{E_{\rho_m}[v_1] - v_1 \geq E_{\rho_m}[v_1] - k\} \tag{A.32}$$
$$\leq Pr_{\rho_m}\{|E_{\rho_m}[v_1] - v_1| \geq |E_{\rho_m}[v_1] - k|\} \tag{A.33}$$
$$= Pr_{\rho_m}\{(E_{\rho_m}[v_1] - v_1)^2 \geq (E_{\rho_m}[v_1] - k)^2\} \tag{A.34}$$
$$\leq \frac{Var_{\rho_m}(v_1)}{(E_{\rho_m}[v_1] - k)^2} \tag{A.35}$$
$$\leq \frac{B}{(\sqrt{N}\lambda_1^m - k)^2} \tag{A.36}$$

Thus, if $K \subset \{0, ..., N/2\}$ such that $|v_1| \leq k$ for $k \in K$, we deduce

$$2\|\rho_m - \pi\|_V = \sum_{i=0}^{N/2} |\rho_m(0,i) - \pi_i| \tag{A.37}$$
$$\geq \sum_K |\rho_m(0,i) - \pi_i| \tag{A.38}$$
$$\geq \sum_K \pi_i - \sum_K \rho_m(0,i) \tag{A.39}$$
$$\geq 1 - \frac{1}{k^2} - \frac{B}{(\sqrt{N}\lambda_1^m - k)^2} \tag{A.40}$$

Choose $k = d\sqrt{N}\lambda_1^m$ to obtain

$$2\|\rho_m - \pi\|_V \geq 1 - \left[\frac{1}{d^2} + \frac{B}{(1-d)^2}\right] N\lambda_1^{-2m} \tag{A.41}$$
$$\geq 1 - be^{4c}, \tag{A.42}$$

which proves theorem 5.2.

References.

REFERENCES

[1] N. Akhiezer, *The Classical Moment Problem*, Hafner, 1965.

[2] A. Baronchelli, C. Castellano, and R. Pastor-Satorras, *Voter models on weighted networks*, Phys. Rev. E, 83 (2011), p. 066117.

[3] A. Baronchelli, M. Felici, V. Loreto, E. Caglioti, and L. Steels, *Sharp transition towards shared vocabularies in multi-agent systems*, J. Stat. Mech.: Theory Exp., (2006), p. P06014, http://stacks.iop.org/1742-5468/2006/i=06/a=P06014.

[4] P. Bremaud, *Markov chain models*, Springer-Verlag, 2005.

[5] F. Caccioli, L. Dall'Asta, T. Galla, and T. Rogers, *Voter models with conserved dynamics*, Phys. Rev. E, 87 (2013), p. 052114.

[6] P. Deift, *Orthogonal Polynomials and Random Matrices*, Courant Lecture Notes 3, 2000.

[7] P. Diaconis and M. Shahshahani, *Time to reach stationarity in the bernoulli–laplace diffusion model*, SIAM Journal on Mathematical Analysis, 18 (1987), pp. 208–218.

[8] L. Flatto, A. M. Odlyzko, and D. Wales, *Random shuffles and group representations*, The Annals of Probability, (1985), pp. 154–178.

[9] B. Friedman, *A simple urn model*, CPAM, 2 (1949), pp. 59–70.

[10] S. Karlin and J. McGregor, *Random walks*, Illinois J. Math., 3 (1959).

[11] N. Masuda, *Voter models with contrarian agents*, Phys. Rev. E, 88 (2013), p. 052803.

[12] N. Masuda, N. Gilbert, and S. Redner, *Heterogeneous voter models*, Phys. Rev. E, 82 (2010), p. 010103(R).
[13] P. Olver, *Equivalence, Invariants and Symmetries*, North Holland, New York, 2005.
[14] W. Pickering, *Solution of Urn Models by Generating Functions with Applications to Social, Physical, Biological, and Network Sciences*, PhD thesis, Dept. Math. Sci., Rensselaer Polytechnic Institute, Troy, NY, 2016.
[15] W. Pickering and C. Lim, *Solution to urn models of pairwise interaction with application to social, physical, and biological sciences*, Phys. Rev. E, 96(1).
[16] W. Pickering and C. Lim, *Solution of the voter model by spectral analysis*, Phys. Rev. E, 91 (2015), p. 012812.
[17] T. Rogers and T. Gross, *Consensus time and conformity in the adaptive voter model*, Phys. Rev. E, 88 (2013), p. 030102(R).
[18] V. Sood and S. Redner, *Voter model on heterogeneous graphs*, Phys. Rev. Lett., 94 (2005), p. 178701.
[19] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. Szymanski, *Social consensus through the influence of committed minorities*, Phys. Rev. E, 84 (2011), p. 011130.
[20] W. Zhang, C. Lim, S. Sreenivasan, J. Xie, B. Szymanski, and G. Korniss, *Social influencing and associated random walk models: Asymptotic consensus times on the complete graph*, Chaos, 21 (2011), p. 025115.