Journal of the Operations Research
Society of Japan
Vol. 26, No. 2, June 1983

AN OPTIMIZATION PROBLEM IN A FINITE MARKOV CHAIN APPLIED TO SOLID PARTICLE MIXING

Keisuke Uchimura
Yamanashi University

Tadashi Yoshizawa
Yamanashi University

Kenichi Shimura
Yamanashi University

(Received January 20, 1981; Final February 25, 1983)

Abstract Consider a Markov chain with n states and a symmetric tridiagonal transition probability matrix P. Let C be a class of vectors $\{p(0)\}$ such that half the elements of $p(0)$ are zero and another half are 1. The problem is to find the vector $p(0)$ in C such that the value $||p(0)P^N - p(\infty)||$ is the minimum for any large nonnegative integer N, where $p(\infty) = (1/2, 1/2, \ldots, 1/2)$ and $p(\infty)$ is called a stationary concentration vector. The solution is stated as follows. The optimal vectors in a sense defined in this paper are of the form $(p_0, p_1, \ldots, p_{n-1})$ which satisfy the following conditions: $p_k = p_{3-k}$, $p_k = p_{4g+k}$ for $0 \leqq k \leqq 3$ and $0 \leqq g \leqq 2^{m-2}-1$, where $n = 2^m$ ($m \geqq 2$). The optimal vectors can be represented as repetition of 0110 or 1001.

This problem was related to models for mixing process of particulate solids. The optimal solutions have been conjectured through the computer simulation conducted by Akao et al. The objective of this paper is to support the mathematical background for their experimental results and to show an interesting property of Markov chains.

1. Introduction

The problem formulated and solved in this paper is to find an optimal initial distribution for Markov chains with some specific transition probability matrix. An initial distribution is said to be optimal if the speed of convergence from the initial distribution to the stationary one is the maximal.

This problem was first suggested by Akao, who has proposed a significant approach to analyse mixing processes of solid particles.

Mixing process in a conventional solids mixer is governed mainly by two basic mechanisms, convection and diffusion. In fact, fairly regular deterministic bulk flow of particles and very irregular stochastic movement of individual particles will be easily observed typically if the two kinds of particles are mixed by a horizontal rotating drum mixer, which is probably one of the simplest types of conventional solids mixer used in various industries. However, real motions are extremely complicated and depend on the shape, density and other physical characteristic of particles as well as types of mixers. Mixing is one of the difficult problems to analyse in chemical engineering.

Akao et al. [1,4] regard mixing process as a synthesized one with two independent operational units which correspond to convention and diffusion. They have proposed a probabilistic branching model for diffusive unit, and studied for convection unit optimal constitution of strata fed to the following diffusive unit. The problems of constitution of strata may be observed, for examples, at strata of cokes and ores in blast furnace in iron-works. If we find the optimal constitution of strata, we can considerably shorten the time required for the mixing.

Numerical and experimental simulation studies on mixing of particulate solids have been reported in Akao et al. [1, 3, 4]. The details of further results and of their applicability to the design of practical mixers will be published in Akao et al. [2]. Concerning the problem solved here, the optimal solutions has been conjectured through their extensive computer simulation [2, 3].

The objective of the present paper is to support the mathematical background for their experimental results. We can regard the constitution of strata as initial distributions for Markov chains. Hence if we can find the optimal initial distribution, we can get the optimal constitution of strata accordingly.

In the second section the probability branching model mixer and the structure of the transition probability matrix will be described on the basis of a Markov chain. The optimization problem to be solved here will be formulated in the third section, where eigenvalues and eigenvectors of the transition probability matrix will be also derived, showing a numerical example. In the fourth section, an algebraic proof of the theorem concerning the optimality will be given. In the fifth section, several modifications of the Markov chain model will be discussed. The approach to this problem will be interesting for readers of this journal since applications of Markov chains are broad in the field of Operation Research.

2. A Probabilistic Branching Model and Feeding Vectors

Akao has proposed an ideally synthesized model both of convective and diffusive mixing processes as shown in Fig. 1, which is a revised version of Fig. 7 in [4]. The upper parts and the central parts of the mixer represent stratified feeders which play a role of convection and a probabilistic branching model mixer which realizes diffusion, respectively. The lower parts are simply receivers of the mixture.

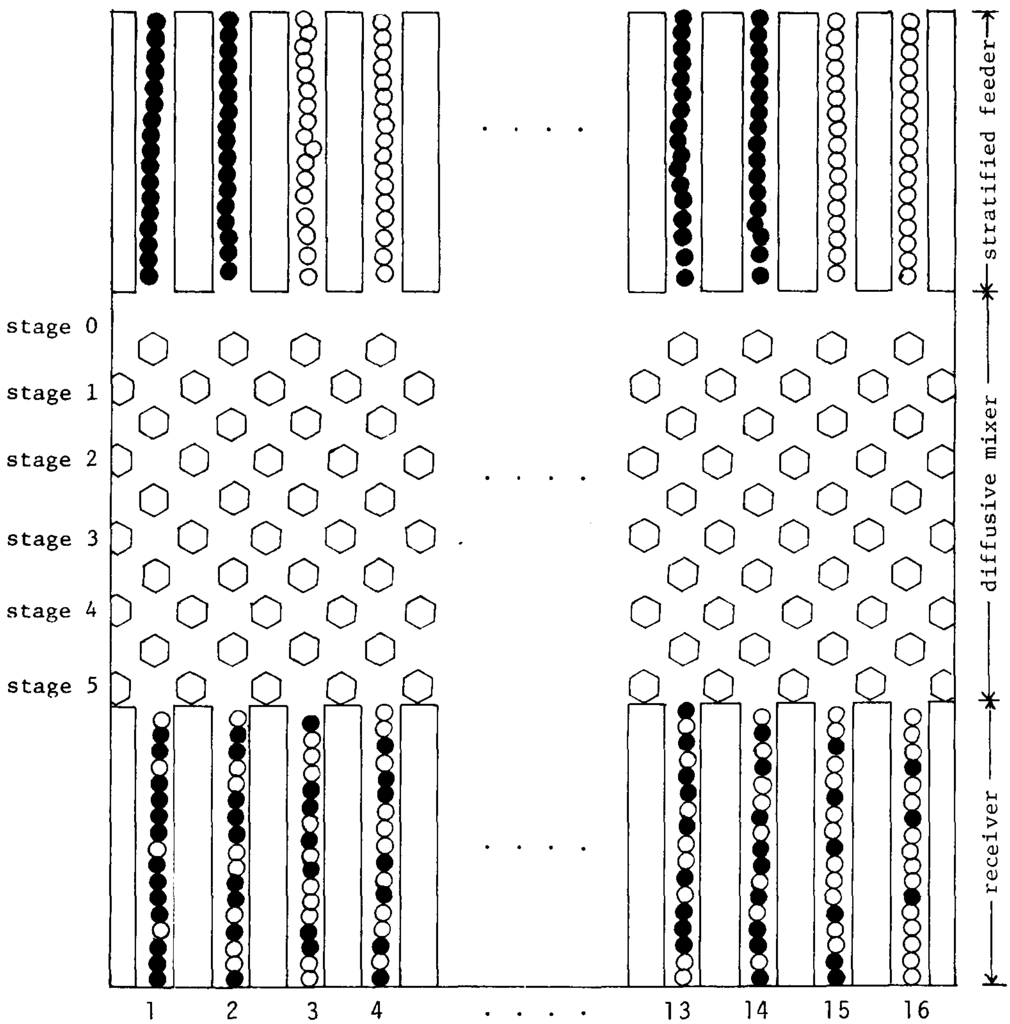

Fig. 1 A stratified feeding mixer and a probabilistic branching model mixer.

This paper deals with the mixture of only two kinds of homogeneous solid particles, whose total concentration is fixed to be one to one. Two kinds of particles are classified into 0 (white) and 1 (black) here for convenience. Each feeder supplies particles either of type 0 or 1, so that the combination of feeders is represented by a vector with the elements of 0 and 1. The vectors are named "feeding vectors". In Fig. 1, the feeding vector is expressed as a row vector (1,1,0,0,1,1,0,0,1,1,0,0,1,1,0,0).

The probabilistic branching model mixer depicted in Fig. 1 is an inclined board with rows of hexagonal blocks set apart by a small distance. Particles with a diameter of slightly less than the distance between the blocks are dropped from the feeders. They are allowed to run down between the hexagonal blocks and are collected in the receivers at the bottom of the board. On their way down, particles are assumed to be independent of each other, that is, no interference exists. The rows numbered by every other rows in Fig. 1 are called stages, at which the locations between hexagonal blocks are called cells. The number of cells is denoted by n. In Fig. 1, each stage has 16 cells.

Motions of particles in the probabilistic branching model in [4] was considered a Markov chain as follows: the cells at each stage constitute a state space; the transition probability p_{ij} from the state i to the state j is the probability that a particle which is located on the i-th cell at any stage moves to the j-th cell at the next stage; the transition probability matrix with the elements p_{ij}'s is of the form

$$(2.1) \qquad P = \begin{pmatrix} 3/4 & 1/4 & & & 0 \\ 1/4 & 1/2 & 1/4 & & \\ & \ddots & \ddots & \ddots & \\ & & 1/4 & 1/2 & 1/4 \\ 0 & & & 1/4 & 3/4 \end{pmatrix},$$

where P is a real symmetric tridiagonal matrix. The quantities of p_{ij}'s can be evaluated under the assumptions that a particle at any row branches with equal probability of 1/2 except at the walls, and both side walls are reflective.

Let us denote a row vector of an initial distribution by $\Pi(0)$ in the case that only one particle is dropped. The element $\Pi_i(0)$ of $\Pi(0)$ means the probability that the particle is supplied from the i-th feeder (i=1, 2, ..., n). It is known that the distribution $\Pi(N)$ at the stage N is given by

(2.2) $\Pi(N) = \Pi(0)P^N$

and that the stationary distribution $\Pi(\infty)$ exists and is equal to $(1/n, 1/n, 1/n)$. Note that the element $\Pi_i(N)$ of $\Pi(N)$ represents the probability that the dropping particle passes the i-th cell at stage N.

Suppose that many particles are supplied continuously and that the number of one kind of particles, e.g. black ones, follows the initial distribution $\Pi(0)$. The quantity $(n/2)\Pi(N)$ can be interpreted as the expected concentration vector whose element represents the concentration ratio of black particles at each cell after N stages.

A feeding vector with 0-1 elements is denoted by $p(0)$. Since $p(0)$ is nothing but an initial concentration vector, $p(N)$ which is defined by

(2.3) $p(N) = p(0)P^N$,

is equivalent to the expected concentration vector mentioned above, i.e.

(2.4) $p(N) = (n/2)\Pi(N)$.

Then, we obtain

(2.5) $$p(\infty) = (n/2)\Pi(\infty) = (1/2, 1/2, \dots, 1/2)$$

as a stationary concentration vector, when the number of stages N approaches to infinity. The mixture with this stationary concentration vector is called "complete".

Akao et al. [2] have shown that the speed of convergence of $P(N)$ to $p(\infty)$ greatly depends on the feeding vectors. Moreover, they have conjectured the feeding vectors which are optimal in a sense. For example, the feeding vector (1,0,0,1,1,0,0,1,1,0,0,1,1,0,0,1) is conjectured to be optimal for the case n=16 in Fig. 1. Note that the distinction of 0 and 1 in feeding vectors is not essential, since the same discussion as above is possible for the concentration of white particles instead of black ones.

3. Statement of the Problem and an example

Consider a finite Markov chain with n states and with the transition probability matrix P of the form

$$(3.1)\qquad P = \begin{bmatrix} p+q & q & & & & \\ q & p & q & & 0 & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ & 0 & & q & p & q \\ & & & & q & p+q \end{bmatrix},$$

where P is a real symmetric tridiagonal matrix of size $n\times n$, and

$$(3.2)\qquad 1 > p \geqq 1/2\ ,$$

and

$$(3.3)\qquad p + 2q = 1.$$

The case in which p is 1/2 is reduced to the model with the transition probability matrix (2.1) described in the previous section, while the other cases in which p is greater than 1/2 may be physically interpreted as models in which some interferences among particles occur. If the condition (3.2) is satisfied, P is positive definite and has the distinct eigenvalues as will be shown in Appendix A. The condition (3.3) shows that the sum of the elements of each row in P is equal to 1.

The problem to be solved here is to find a feeding vector $p(0)$ such that the expected concentration vector $p(N)$ defined by (2.3) converges to the stationary concentration vector $p(\infty)$ as fast as possible. The speed of convergence may be measured with the smallest value of N that satisfied the condition;

$$(3.4)\qquad \| p(N) - p(\infty)\|_E < \varepsilon$$

for a sufficiently small ε, where $\|.\|_E$ denotes Euclidean norm about vectors.

Now, we denote the eigenvalues and the corresponding eigenvectors of the transition probability matrix P by λ_j and $\bar{u}_j$ (j=0, 1, ..., n-1) respectively, i.e.

$$(3.5)\qquad P\bar{u}_j = \lambda_j\bar{u}_j, \qquad j=0,\ 1,\ \ldots,\ n-1,$$

where λ_j's are arranged as

$$(3.6)\qquad \lambda_0 > \lambda_1 > \ldots > \lambda_{n-1} > 0\ ,$$

and $\bar{u}_j$'s are assumed to be normalized as

$$(3.7)\qquad {}^t\bar{u}_j\bar{u}_j = 1, \qquad j=0,\ 1,\ \ldots,\ n-1,$$

so that they are orthonormal. Note that $\bar{u}_i$ is a column vector, while $p(N)$ is a row vector, and that t means transposition of a vector or a matrix, hereafter.

It is well known that the maximum eigenvalue of the probability matrix P is 1, and that P is decomposable to the form

$$(3.8)\qquad P = \sum_{j=0}^{n-1} \lambda_j \bar{u}_j {}^t\bar{u}_j .$$

From the above facts and the orthonormality of the $\bar{u}_j$'s, we can easily obtain

$$(3.9)\qquad \begin{aligned} P^N &= \sum_{j=0}^{n-1} \lambda_j^N \bar{u}_j {}^t\bar{u}_j \\ &= \bar{u}_0 {}^t\bar{u}_0 + \sum_{j=1}^{n-1} \lambda_j^N \bar{u}_j {}^t\bar{u}_j . \end{aligned}$$

Therefore, the expected concentration vector $p(N)$ of (2.3) is expressed as

$$(3.10)\qquad \begin{aligned} p(N) &= p(0)P^N = \sum_{j=0}^{n-1} \lambda_j^N (p(0)\bar{u}_j) {}^t\bar{u}_j \\ &= a_0 {}^t\bar{u}_0 + \sum_{j=1}^{n-1} \lambda_j^N a_j {}^t\bar{u}_j , \end{aligned}$$

where a_j is the inner product of $p(0)$ and $\bar{u}_j$, i.e.,

$$(3.11)\qquad a_j = p(0)\bar{u}_j, \qquad 0 \leqq j \leqq n-1 .$$

Since the maximal eigenvalue λ_0 is 1 and the other λ_j's are less than 1, we obtain

$$(3.12)\qquad p(\infty) = a_0 {}^t\bar{u}_0 .$$

From (3.10) and (3.12), inequality (3.4) can be written as

$$(3.13)\qquad \| p(N) - p(\infty) \|_E = \| \sum_{j=1}^{n-1} \lambda_j^N a_j {}^t\bar{u}_j \|_E = \Big(\sum_{j=1}^{n-1} \lambda_j^{2N} a_j^2 \Big)^{1/2} < \varepsilon .$$

Then, let us define the quantity $\sigma(N)$ as

$$(3.14)\qquad \sigma(N) = \| \sum_{j=1}^{n-1} \lambda_j^N a_j {}^t\bar{u}_j \|_E / n^{1/2} .$$

It is easily observed that $\sigma(N)$ will be dominated by $\lambda_1^{2N} a_1^2$ for a sufficiently large N if a_1 is not equal to zero. $\lambda_k^{2N} a_k^2$ is said to be the principal term of $\sigma(N)$ if a_k is the first nonzero factor in (3.14), that is

$$a_1 = a_2 = \dots = a_{k-1} = 0, \quad \text{and} \quad a_k \neq 0.$$

Consider a class C of feeding vectors $\{p(0)\}$ such that half the elements of $p(0)$ are zero and another half are 1. We order the vectors in C as follows. For any p_1 and p_2 in C, let

(3.15) $\quad p_1\bar{u}_1 = \dots = p_1\bar{u}_{k-1} = 0, \quad p_1\bar{u}_k \neq 0$ and

(3.16) $\quad p_2\bar{u}_1 = \dots = p_2\bar{u}_{s-1} = 0, \quad p_2\bar{u}_s \neq 0.$ Then,

(3.17) $\quad p_1 \leq p_2$ if and only if $k > s$ or both $k = s$ and $|p_1\bar{u}_k| \leq |p_2\bar{u}_k|$ hold.

The minimal elements are called the optimal feeding vectors. It is clear that if $p(0)$ is an optimal feeding vector, the smallest integer N satisfying the condition (3.4) with $p(0)$ is the smallest of all integers with vectors in C.

Problem Find the optimal feeding vectors in C.

Note that the feeding vector $p(0)$ is regard to be proportioned to an initial distribution $\Pi(0)$ as mentioned in the previous section, so that this problem can be stated concerning initial distribution instead of feeding vectors.

Now, let us show the eigenvalues and eigenvectors of the transition probability matrix P of (3.1). The details of their derivation will be described in Appendix A. The eigenvalues of P are given by

(3.18) $\quad \lambda_j = p + y_j q, \qquad 0 \leq j \leq n-1,$

where

(3.19) $\quad y_j = 2\cos(j\pi/n), \qquad 0 \leq j \leq n-1.$

The corresponding eigenvectors u_j, whose elements are denoted by $u_{j,k}$ $(0 \leq k \leq n-1)$, are expressed as

(3.20) $\quad u_{j,k} = \cos((2k+1)j\pi/2n), \quad j=0, 1, \dots, n-1.$

Note that

(3.21) $\quad \bar{u}_0 = u_0/n^{1/2}, \quad \bar{u}_j = u_j/(n/2)^{1/2}, \quad j \geq 1.$

It is worthy of notice that the eigenvectors do not depend on p and q. Since y_j's are distinct and

$$|y_j| \leqq 2, \quad j=0,\ 1,\ \ldots,\ n-1,$$

the eigenvalues of P are distinct and positive if p and q satisfy (3.2) and (3.3). Note that λ_0 is equal to 1 as well known, and that

$$u_0 = {}^t(1,\ 1,\ \ldots,\ 1).$$

Thus, we know from (3.11),

$$a_0 = n^{1/2}/2$$

and

$$p(\infty) = a_0{}^t u_0/({}^t u_0 u_0)^{1/2} = (1/2,\ \ldots,\ 1/2),$$

for any $p(0)$ which belongs to the class restricted in the problem.

The number of entries of the class $\{p(0)\}$ is finite for any fixed n, so that the optimal feeding vectors can be determined if all entries are surveyed. Let us show an example below. Remark that the eigenvectors in Table 1 are not normal.

Example 1. $n = 16$

Table 1. Eigen values and vectors of the stochastic matrix P_{16}

ele- ments	eigen vectors 0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	1.	1.	.98	.96	.92	.88	.83	.77	.71	.63	.56	.47	.38	.29	.20	.10
1	1.	.96	.83	.63	.38	.10	-.20	-.47	-.71	-.88	-.98	-1.	-.92	-.77	-.56	-.29
2	1.	.88	.56	.10	-.38	-.77	-.98	-.96	-.71	-.29	.20	.63	.92	1.	.83	.47
3	1.	.77	.20	-.47	-.92	-.96	-.56	.10	.71	1.	.83	.29	-.38	-.88	-.98	-.63
4	1.	.63	-.20	-.88	-.92	-.29	.56	1.	.71	-.10	-.83	-.96	-.38	.47	.98	.77
5	1.	.47	-.56	-1.	-.38	.63	.98	.29	-.71	-.96	-.20	.77	.92	.10	-.83	-.88
6	1.	.29	-.83	-.77	.38	1.	.20	-.88	-.71	.47	.98	.10	-.92	-.63	.56	.96
7	1.	.10	-.98	-.29	.92	.47	-.83	-.63	.71	.77	-.56	-.88	.38	.96	-.20	-1.
8	1.	-.10	-.98	.29	.92	-.47	-.83	.63	.71	-.77	-.56	.88	.38	-.96	-.20	1.
9	1.	-.29	-.83	.77	.38	-1.	.20	.88	-.71	-.47	.98	-.10	-.92	.63	.56	-.96
10	1.	-.47	-.56	1.	-.38	-.63	.98	-.29	-.71	.96	-.20	-.77	.92	-.10	-.83	.88
11	1.	-.63	-.20	.88	-.92	.29	.56	-1.	.71	.10	-.83	.96	-.38	-.47	.98	-.77
12	1.	-.77	.20	.47	-.92	.96	-.56	-.10	.71	-1.	.83	-.29	-.38	.88	-.98	.63
13	1.	-.88	.56	-.10	-.38	.77	-.98	.96	-.71	.29	.20	-.63	.92	-1.	.83	-.47
14	1.	-.96	.83	-.63	.38	-.10	-.20	.47	-.71	.88	-.98	1.	-.92	.77	-.56	.29
15	1.	-1.	.98	-.96	.92	-.88	.83	-.77	.71	-.63	.56	-.47	.38	-.29	.20	-.10
eigen values	1.	.99	.96	.92	.85	.78	.69	.60	.50	.40	.31	.22	.15	.08	.04	.01

We have computed the integer k such that a_k is the first nonzero factor in (3.14) and the value a_k for fifteen feeding vectors of size 16. Further we have evaluated the minimal integer N_0 that satisfies the inequality

$$\sigma^2(N_0) \leqq 1/4000, \quad (3.22)$$

for each feeding vector by direct computation.

Besides we have determined another integer M for each feeding vector as follows. Let $\lambda_k^{2N} a_k^2$ be the principal term of $\sigma(N)$ for the vector. We have determined the minimal integer M that satisfies the inequality

$$\lambda_k^{2M} a_k^2/16 \leqq 1/4000. \quad (3.23)$$

The result is depicted in the following table.

Table 2. Speed of Convergence

	patterns	k	a_k	N_0	M
A	1111111100000000	1	1.80	348	348
B	1001100110011001	8	2.00	5	5
C	1111000011110000	1	0.75	256	256
D	1111000000001111	2	1.81	87	87
E	1100110011001100	1	0.36	180	180
F	1100110000110011	2	0.75	64	64
G	1100001111000011	4	1.85	22	22
H	1010101010101010	1	0.18	107	107
I	1010101001010101	2	0.36	45	45
J	1010100110010101	2	0.25	36	36
K	1010011001100101	2	0.11	17	14
L	1001010110101001	2	-0.25	36	36
M	1001011001101001	2	-0.15	23	23
N	1001101001011001	2	0.11	17	14
O	1010010110100101	4	0.77	16	16

In this table we can see that $\sigma^2(N)$ is dominated by the principal term $\lambda_k^{2N} a_k^2$ for any large N.

Here we explain the reason for using the value 1/4000 in (3.22) and (3.23).

Akao et al. compared the speed of convergence among the fifteen feeding vectors by computer simulation. Although $\sigma(N)$ becomes zero in the complete mixed state, the simulation results are acompanied by the error of the simple random sampling. The value $\sigma_r^{\ 2}$ of the error is evaluated to be

$$\sigma_r^{\ 2} = p(1-p)/k$$

based on the binomial distribution, where k is the number of particles fed at each cell (See, [2]). The criterion of convergence can be determined as

$$\sigma^2(N) \leqq \sigma_r^{\ 2}/10$$

from the convention which is practically employed in the field of sampling. In their simulation 100 particles were fed.

In order to compare our result with theirs, we chose the value 1/4000 and evaluated the integers N_0 and M.

In Table 2 we see the fact that the speed of convergence of the pattern B is the fastest, which accords with the result of the computer simulation stated in [2].

4. Optimal Solutions in the case that $n = 2^m$

In this section, the theorem which gives the optimal solutions of the problem will be proved, where $n = 2^m$.

Theorem. Where $n = 2^m$ $(m \geqq 2)$, the optimal feeding vectors of the problem stated in the previous section are the feeding vectors of the form $(p_0, p_1, \ldots, p_{n-1})$ which satisfy the following condition c_{m-2};

$$c_{m-2} \;:\; p_k = p_{4-1-k},$$
$$p_k = p_{4\times q+k},$$
$$0 \leqq k \leqq 3, \quad 0 \leqq g \leqq 2^{m-2}-1.$$

Representing the optimal vector $(p_0, p_1, \ldots, p_{n-1})$ as a string $p_0p_1 \ldots\ldots p_{n-1}$, we have

$$p_0p_1 \cdots p_{n-1} = \alpha_{m-2} ,$$

where $\alpha_i = \alpha_{i-1}*\alpha_{i-1}$, and $\alpha_0 = 0110$ (or 1001) (the symbol * denotes a concatenation product of two strings).

To prove the theorem we need some preparation. Recall that from (3.20) we have

$$u_{j,k} = \cos\,((2k+1)j\pi/2n), \quad 0 \leqq j,k \leqq n-1.$$

Let $\zeta = \exp\,(2\pi i/4n)$. Clearly ζ is a primitive $4n$-th root of unity. Then we have

$$2u_{j,k} = \zeta^{j(2k+1)} + \zeta^{-j(2k+1)} , \quad 0 \leqq k \leqq n-1 ,$$

since $\zeta = \cos(2\pi/4n) + i\sin(2\pi/4n)$. Note that

$$\zeta^{2^{m+2}} = 1 \quad\text{and}\quad \zeta^{2^{m+1}} = -1.$$

First we show a condition for $p(0)$ to make the equality $p(0)u_1 = 0$ valid.

Note that $2u_1 = {}^t(\zeta+\zeta^{-1}, \zeta^3+\zeta^{-3}, \ldots, \zeta^{2n-1}+\zeta^{-(2n-1)})$ and each element of $p(0)$ is 1 or 0.

Lemma 1. The equality,

$$\sum_{k=0}^{2^m-1} c_k(\zeta^{2k+1} + \zeta^{-(2k+1)}) = 0, \tag{4.1}$$

where each c_k is in the rational number field Q, is valid if and only if

$$c_k = c_{2^m-1-k} \quad \text{for all } k = 0, 1, \ldots, 2^m-1.$$

Proof: Let $\phi_{2^{m+2}}(x)$ be the 2^{m+2}-th cyclotomic polynomial. It is known that

$$\phi_{2^{m+2}}(x) = \prod_{d|2^{m+2}} (x^d-1)^{\mu(2^{m+2}/d)}$$

where μ is the Moebius function (See Van der Waerden [6]). Hence

$$\phi_{2^{m+2}}(x) = x^{2^{m+1}} + 1 .$$

It is known that any cyclotomic polynomial is irreducible over Q. Therefore $x^{2^{m+1}}+1$ is irreducible over Q. This tells us that the polynomial $x^{2^{m+1}}+1$ is the minimal polynomial over Q for ζ. Multiply the eq. (4.1) by $\zeta^{2^{m+1}-1}$
We get

$$\sum_{k=0}^{2^m-1} c_k(\zeta^{2k+2^{m+1}} + \zeta^{2^{m+1}-2k-2}) = 0.$$

Since $\zeta^{2^{m+1}} = -1$, we have

$$(-c_0+c_{2^m-1}) + \zeta^2(-c_1+c_{2^m-2}) + \ldots + \zeta^{2^{m+1}-4}(-c_{2^m-2}+c_1) + \zeta^{2^{m+1}-2}(-c_{2^m-1}+c_0) = 0.$$

Therefore we have

$$c_k = c_{2^m-1-k}, \qquad (0 \leq k \leq 2^m-1),$$

since $x^{2^{m+1}}+1$ is the minimal polynomial for ζ. Q. E. D.

Next we show the condition of $p(0)$ for $p(0)u_j = 0$ when j is an odd integer.

Recall that

$$2u_j = {}^t(\zeta^j+\zeta^{-j},\ \zeta^{3j}+\zeta^{-3j},\ \ldots,\ \zeta^{j(2n-1)}+\zeta^{-j(2n-1)}).$$

Corollary 1. Let j be an odd integer. Then

$$\sum_{k=0}^{2^m-1} c_k(\zeta^{j(2k+1)} + \zeta^{-j(2k+1)}) = 0,$$

if and only if $c_k = c_{2^m-1-k}$ for all k, $(0 \leq k \leq 2^m-1)$.

The proof follows from the fact that ζ^j is a primitive $4n$-th root of unity for any odd integer j.

We classify the set of eigenvectors $\{u_j | j=1, 2, \ldots, 2^m-1\}$ as follows. Let $V_r = \{u_j | j$ is divisible by 2^r but not divisible by $2^{r+1}\}$. Note that $0 \leq r \leq m-1$.

Lemma 2. The following property is valid for any eigenvector $u_j \in V_r$ whenever $1 \leq r$.

$$u_{j,k} = u_{j,2^{m-r+1}-1-k}, \qquad (0 \leq k \leq 2^{m-r+1}-1),$$

$$u_{j,k} = u_{j,2^{m-r+1}\times g+k}, \qquad (0 \leq g \leq 2^{r-1}-1).$$

Proof: First we show the lemma holds for u_{2^r}. Clearly,

$$\begin{aligned} 2u_{2^r,k} &= \zeta^{2^r(2k+1)} + \zeta^{-2^r(2k+1)} \\ &= \zeta^{-2^r(2^{m-r+2}-2k-1)} + \zeta^{2^r(2^{m-r+2}-2k-1)} \\ &= 2u_{2^r,2^{m-r+1}-1-k}, \end{aligned}$$

since $\zeta^{2^{m+2}} = 1$.

Also for any g we have

$$\begin{aligned} 2u_{2^r,k} &= \zeta^{2^r(2k+1)} + \zeta^{-2^r(2k+1)} \\ &= \zeta^{2^r(2^{m-r+2}\times g+2k+1)} + \zeta^{-2^r(2^{m-r+2}\times g+2k+1)} \\ &= u_{2^r,2^{m-r+1}\times g+k}. \end{aligned}$$

By the similar method, we can prove that the lemma is valid for any $u_j \in V_r$, since we use only the property; $\zeta^{2^{m+2}} = 1$ to prove the lemma for u_{2^r}. Q.E.D.

This situation can be seen clearly in Table 1.

We shall find a necessary and sufficient condition for $p(0)$ to make the equality $p(0)u = 0$ valid for every u in $V_0 \cup V_1 \cup \cdots \cup V_r$. Let

$$p(0) = (p_0, p_1, \ldots, p_{2^m-1}).$$

We denote by C_r the following condition;

$$p_k = p_{2^{m-r}-1-k} \qquad (0 \leqq k \leqq 2^{m-r}-1),$$

and

$$p_k = p_{2^{m-r}\times g+k}, \qquad (0 \leqq g \leqq 2^r-1).$$

Lemma 3. The equality $p(0)u = 0$ is valid for every u in $V_0 \cup V_1 \cup \cdots$ $\cdots \cup V_r$, if and only if $p(0)$ satisfies the condition C_r.

Proof: By induction. If $r = 0$, we can prove the assertion by Lemma 1 and Collorary 1.

Assume the lemma proved for $r-1$. Then

$$p_k = p_{2^{m-r+1}-1-k},$$

(4.2) $$p_k = p_{2^{m-r+1}\times g+k},$$

$$(0 \leqq k \leqq 2^{m-r+1}-1), \qquad (0 \leqq g \leqq 2^{r-1}-1).$$

From Lemma 2, for any $u_j \in V_r$, we find

(4.3) $$u_{j,k} = u_{j,2^{m-r+1}-1-k}, \qquad u_{j,k} = u_{j,2^{m-r+1}\times g+k},$$

where $j = 2^r \times q$ and q is an odd integer. We may assume without loss of generality that $q = 1$. Assume that $p(0)u_{2^r} = 0$. Then we have

$$\sum_{k=0}^{2^{m-r}-1} p_k(\zeta^{2^r(2k+1)} + \zeta^{-2^r(2k+1)}) = 0.$$

It is obvious that ζ^{2^r} is a primitive 2^{m-r+2}-th root of unity. Hence from Lemma 1 and (4.2) we have

$$p_k = p_{2^{m-r}-1-k},$$

$$p_k = p_{2^{m-r}\times g+k},$$

$$(0 \leq k \leq 2^{m-r}-1), \qquad (0 \leq g \leq 2^r-1).$$

This is the condition C_r. Q. E. D.

Proof of Theorem: As stated in Lemma 3, the following properties are equivalent to each other:

(1) $p(0)u_i = 0$ for all u_i $(1 \leq i \leq 2^r-1)$.

(2) $p(0)u_{i_j} = 0$ for every u_{i_j} $(1 \leq j \leq r)$, where u_{i_j} is any element of V_{j-1} $(1 \leq j \leq m)$ and $1 \leq r \leq m$.

If $p(0)u_{i_j} = 0$ for every u_{i_j} $(1 \leq j \leq m)$, $p(0) = (0,0, \ldots, 0)$ or $(1,1, \ldots, 1)$. Since half the elements of $p(0)$ are zero and another half are 1, we have the optimum feeding vector $p(0)$ if $p(0)u_{i_j} = 0$ $(1 \leq j \leq m-1)$ and $p(0)u_{i_m} \neq 0$. Hence it follows from Lemma 3 that the optimum feeding vector $p(0)$ is equal to $p_1 = (0,1,1,0,0,1,1,0, \ldots, 0,1,1,0)$ or $p_2 = (1,0,0,1,1,0,0,1, \ldots, 1,0,0,1)$.

Q. E. D.

5. Discussion

In this section we consider several modification of the previous Markov chain model.

(a) Classes of feeding vectors

In Problem we assume that half the elements of $p(0)$ are zero and another half are 1. However we can apply our theory to the case where the total concentration ratio is not equal to 1/2.

For instance we consider the following case: $n = 16$ and the total concentration ratio is equal to 1/4. By similar analysis we can show that the optimal feeding vectors are included in the following vectors.

$F1$: (1,0,0,0,0,0,0,1,1,0,0,0,0,0,0,1),
$F2$: (0,1,0,0,0,0,1,0,0,1,0,0,0,0,1,0),
$F3$: (0,0,1,0,0,1,0,0,0,0,1,0,0,1,0,0),
$F4$: (0,0,0,1,1,0,0,0,0,0,0,1,1,0,0,0).

Clearly $a_i (1 \leqq i \leqq 3)$ is zero for any Fj $(1 \leqq j \leqq 4)$. But a_4 is not zero for any of these feeding vectors.

$$a_4 = 1.30 \quad \text{for} \quad F1,$$
$$a_4 = 0.54 \quad \text{for} \quad F2,$$
$$a_4 = -0.54 \quad \text{for} \quad F3,$$
$$a_4 = -1.30 \quad \text{for} \quad F4.$$

Thus $F2$ and $F3$ are the optimal feeding vectors.

(b) $n \neq 2^m$

In case $n \neq 2^m$ the similar theorem as in case $n = 2^m$ has not yet been obtained. But we can find the eigenvectors. The number of vectors in C is finite for any fixed n, so that the optimal feeding vectors can be determined if all vectors are surveyed.

(c) A cubic model

A cubic model of a stratified feeding mixer and a probabilistic branching model mixer may be considered.

The characteristic equations for the transition probability matrix of this model can be represented by the same equation as $S_n(y)$ in Appendix A, whose argument y is a matrix. Our theory can be applied also to this model and it shows the following pattern to be optimal for $n = 16$.

0110
1001
1001
0110

The details of this model will be reported elsewhere.

(d) The case that interference among particles exists.

Some kind of interference may be modeled by taking p greater that 1/2. In section 3, we showed that eigenvectors do not depend on the values of p and q. Then, our result for optimality still follows. However, real mixing processes show more complex interference, and so we need further steps in this research.

Acknowledgment

The authors wish to thank Professor Y. Akao and Mr. H. Shindo for their helpful suggestions and critical reading of the paper.

References

[1] Akao, Y., Yoshizawa, H., Uehara, S. and Noda, T.: Studies on the Mixing Processes Based on an Elementary Model. *Proceedings of the Tohoku Conference of the Society of Chemical Engineers, Japan*, 1970.

[2] Akao, Y., Shindo, H. and Angel, H.: Synthesis of a Mixing System and Its Optimal Feeding Strata. *Journal of the Soeciety of Powder Technology, Japan*, Vol.19, No.3 (1982), 639-645.

[3] Angel, H., Iwanami, K., Suzuki, R., Shindo, H. and Akao, Y.: Synthesis of a Mixing System and Its Optimal Feeding Strata. *Proceeding of the 11th Conference of the Society of Chemical Engineers, Japan*, 1977.

[4] Fan, L. T., Lai, F. S., Akao, Y., Shinoda, K. and Yoshizawa, H.: Numerical and Experimental Simulation Studies on the Mixing of Particulate Solids and the Synthesis of a Mixing System. *Computers and Chemical Engineering*, Vol.2, (1978), 19-32.

[5] Snyder, M. A.: *Chebyshev Methods in Numerical Approximation*. Prentice Hall, N, J., 1966.

[6] Van der Waerden, B. L.: *Modern Algebra*. Springer-Verlarg, Berlin, 1937.

Keisuke UCHIMURA: Department of Computer Science, Faculty of Engineering, Yamanashi University, Takeda 4-3-11, Kofu, Yamanashi, 400, Japan.

Appendix A

Eigenvalues and eigenvectors of the probability matrix P.

$$\det(P-\lambda E) = \begin{vmatrix} p+q-\lambda & q & & & & \\ q & p-\lambda & q & & 0 & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & 0 & & q & p-\lambda & q \\ & & & & q & p+q-\lambda \end{vmatrix}$$

$$= (1-\lambda)q^{n-1} \begin{vmatrix} 1 & 1 & & & & & \\ 1 & (p-\lambda)/q & 1 & & & 0 & \\ 1 & 1 & (p-\lambda)/q & 1 & & & \\ \cdot & & \cdot & \cdot & \cdot & & \\ \cdot & & & \cdot & \cdot & \cdot & \\ \cdot & & & & \cdot & \cdot & \cdot \\ & 0 & & & 1 & (p-\lambda)/q & 1 \\ 1 & & & & & 1 & 1+(p-\lambda)/q \end{vmatrix},$$

where E denotes an $n\times n$ unit matrix.

Let $y = (p-\lambda)/q$.

Let $$S_{n-1}(y) = \begin{vmatrix} 1 & 1 & & & & \\ 1 & y & 1 & & 0 & \\ 1 & 1 & y & 1 & & \\ \cdot & & \cdot & \cdot & \cdot & \\ \cdot & & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot & \cdot \\ & 0 & & 1 & y & 1 \\ 1 & & & & 1 & 1+y \end{vmatrix},$$

where the right side is the determinant of an $n\times n$ matrix. It can be easily seen that

$S_2(y) = y^2 - 1$, $S_3(y) = y^3 - 2y$, and $S_n(y) - yS_{n-1}(y) + S_{n-2}(y) = 0$,

for any $n \geqq 4$. Note that $S_n(y)$ is a polynomial of degree n being odd or even as n is odd or even. It is clear that if n is odd, then $S_n(y)$ is divisible by y and $S_n(y)/y$ is even. It follows that if z is a zero of $S_n(y)$, then $-z$ is also a zero of $S_n(y)$. It is known that $S_n(y)$ is a version of the Chebyshev polynomial of the second kind (See [4]). Hence

$$S_{n-1}(2\cos\theta) = (\sin n\theta)/\sin\theta.$$

Clearly each zero y_j of this polynomial is determined by the element θ such that $\sin n\theta = 0$ and $\sin\theta \neq 0$.

Thus $y_j = 2\cos(j\pi/n)$. By the remark mentioned above, $-y_j$ $(1 \leqq j \leqq n-1)$ is also a zero of $S_{n-1}(y)$. Then we order λ_j's as follows:

$$\lambda_j = p + y_j q, \qquad 1 \leqq j \leqq n-1.$$

Since $(1-\lambda)$ divides $\det(P-\lambda E)$, it follows that $\lambda_0 = 1$. It is clear that

$$\lambda_0 > \lambda_1 > \lambda_2 > \cdots > \lambda_{n-1}.$$

Further if $p \geqq 2q$, then $\lambda_{n-1} > 0$.

Next we show eigenvectors of P. Clearly an eigenvector corresponding to λ_0 is the vector $(1, \ldots, 1)$. For each eigenvalue λ_j $(1 \leqq j \leqq n-1)$, we shall find an eigenvector $(u_{j,0}, \ldots, u_{j,n-1})$. Clearly it satisfies the following condition:

$$\begin{pmatrix} 1-y_j & 1 & & & & 0 \\ 1 & -y_j & 1 & & & \\ & \cdot & \cdot & \cdot & & \\ & & \cdot & \cdot & \cdot & \\ & & & \cdot & \cdot & \cdot \\ 0 & & & 1 & -y_j & 1 \\ & & & & 1 & 1-y_j \end{pmatrix} \begin{pmatrix} u_{j,0} \\ u_{j,1} \\ \cdot \\ \cdot \\ \cdot \\ u_{j,n-2} \\ u_{j,n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 0 \end{pmatrix}$$

Let $\zeta = \exp(2\pi i/4n)$. Clearly $\zeta^{4n}=1$ and $\zeta = \cos(2\pi/4n) + i\sin(2\pi/4n)$. Hence $y_i = 2\cos(j\pi/n) = \zeta^{2j} + \zeta^{-2j}$. Let $u_j = {}^t(\zeta^j+\zeta^{-j}, \ldots, \zeta^{j(2k+1)} + \zeta^{-j(2k+1)}, \ldots, \zeta^{j(2n-1)} + \zeta^{-j(2n-1)})/2$. Then the vector u_j is an eigenvector to λ_j. Indeed.

$$(1-y_j)x_0 + x_1 = (1-(\zeta^{2j}+\zeta^{-2j}))(\zeta^j+\zeta^{-j}) + \zeta^{3j} + \zeta^{-3j} = 0.$$

$$x_{k-1} - y_j x_k + x_{k+1} = \zeta^{j(2k-1)} + \zeta^{-j(2k-1)} - (\zeta^{2j}+\zeta^{-2j})(\zeta^{j(2k-1)}+\zeta^{-j(2k+1)})$$
$$+ \zeta^{j(2k+3)} + \zeta^{-j(2k+3)} = 0,$$

$$x_{n-2} + (1-y_j)x_{n-1} = 0, \quad \text{where} \quad x_k = \zeta^{j(2k+1)} + \zeta^{-j(2k+1)}, \quad 0 \leqq k \leqq n-1.$$

Note that $(u_j, u_k) = 0$ if $j \neq k$

$= n/2$ if $j = k$