A SYSTEMATIC MARTINGALE CONSTRUCTION WITH APPLICATIONS TO PERMUTATION INEQUALITIES

VLADIMIR POZDNYAKOV AND J. MICHAEL STEELE

ABSTRACT. We illustrate a process that constructs martingales with help from matrix products that arise naturally in the theory of sampling without replacement. The usefulness of the new martingales is illustrated by the development of maximal inequalities for permuted sequences of real numbers. Some of these inequalities are new and some are variations of classical inequalities like those introduced by A. Garsia in the study of rearrangement of orthogonal series.

KEYWORDS: construction of martingales, permutation inequalities, Garsia inequality, combinatorial martingales, discrete Brownian bridge.

MATHEMATICS SUBJECT CLASSIFICATION (2000): 60G42, 60C05

1. INTRODUCTION: A MOTIVATING QUESTION

Let $\mathcal{X} = \{x_1, x_2, ..., x_n\}$ be a fixed set of real numbers, and let $X_1, X_2, ..., X_n$ be the successive values of a sample of size n that is drawn sequentially without replacement from the set $\mathcal{X}$. We are concerned here with a systematic process by which one can construct martingales with respect to the sequence of sigma-fields $\mathcal{F}_k = \sigma(X_1, X_2, ..., X_k)$, $1 \le k \le n$, where, as usual, $\sigma(X_1, X_2, ..., X_k)$ is the σ-field generated by the random variables $\{X_1, X_2, ..., X_k\}$. For example, we consider the partial sums

$$S_k = X_1 + X_2 + \cdots + X_k, \quad 1 \le k \le n,$$

and ask: Is there an $\{\mathcal{F}_k : 1 \le k \le n\}$ martingale where the values $\{S_k^2 : 1 \le k \le n\}$ appear in an simple and explicit way? What about the values $\{S_k^3 : 1 \le k \le n\}$, or the partial sums of X_i^2, etc. We show that there is a practical, unified approach to these problems. It faces some limitations due to the burden of algebra, but one can make considerable progress before those burdens become too cumbersome.

We first illustrate the construction with two basic examples. These lead in turn to several martingales whose usefulness we indicate by the derivation of permutation inequalities — both old and new.

2. FIRST EXAMPLE OF THE CONSTRUCTION

To begin, we consider $T_k = X_1^2 + X_2^2 + \cdots + X_k^2$ for $1 \le k \le n$, and we ask for a martingale where T_k (or a deterministic multiple of T_k) appears as a summand. The set $\mathcal{X}$ is known before sampling beings, and the only source of randomness is

the sampling process itself. The totals are known, deterministic, values which we denote by

(1) $$M = S_n \quad \text{and} \quad B = T_n.$$

We first compute the conditional expectation of S_{k+1}^2 given the immediate past. That is, we let $\mathcal{F}_k = \sigma(X_1, X_2, \ldots, X_k)$ denote the σ-field that is generated by $\{X_1, X_2, \ldots, X_k\}$, and we compute

$$
\begin{aligned}
E[S_{k+1}^2|\mathcal{F}_k] &= E[(S_k + X_{k+1})^2|\mathcal{F}_k] \\
&= E[S_k^2 + 2S_k X_{k+1} + X_{k+1}^2|\mathcal{F}_k] \\
&= S_k^2 + 2S_k \frac{M - S_k}{n - k} + \frac{B - T_k}{n - k} \\
&= \frac{n - k - 2}{n - k} S_k^2 + \frac{2M}{n - k} S_k - \frac{1}{n - k} T_k + \frac{B}{n - k}.
\end{aligned}
$$

To organize this information, introduce the vector-column $\boldsymbol{\xi}_k = (S_k^2, S_k, T_k, 1)^\top$, and note that we also have

$$E[S_{k+1}|\mathcal{F}_k] = S_k + \frac{M - S_k}{n - k} = \frac{n - k - 1}{n - k} S_k + \frac{M}{n - k}, \quad \text{and}$$

$$E[T_{k+1}|\mathcal{F}_k] = T_k + \frac{B - T_k}{n - k} = \frac{n - k - 1}{n - k} T_k + \frac{B}{n - k}.$$

These observations can be combined into one matrix equation for the one-step conditional expected values

$$E[\boldsymbol{\xi}_{k+1}|\mathcal{F}_k] = A_{k+1}\boldsymbol{\xi}_k,$$

where here the deterministic 4×4 matrix A_{k+1} is given explicitly by

$$
A_{k+1} = \begin{pmatrix}
\dfrac{n - k - 2}{n - k} & \dfrac{2M}{n - k} & -\dfrac{1}{n - k} & \dfrac{B}{n - k} \\[2mm]
0 & \dfrac{n - k - 1}{n - k} & 0 & \dfrac{M}{n - k} \\[2mm]
0 & 0 & \dfrac{n - k - 1}{n - k} & \dfrac{B}{n - k} \\[2mm]
0 & 0 & 0 & 1
\end{pmatrix}.
$$

The matrices $\{A_k : 1 \leq k \leq n - 2\}$ are invertible and deterministic, so the vector process

(2) $$\boldsymbol{M}_k = A_1^{-1} A_2^{-1} \cdots A_k^{-1} \boldsymbol{\xi}_k$$

is well-defined and adapted to $\{\mathcal{F}_k : 1 \leq k \leq n - 2\}$. To check that it is a vector martingale we only need to note that

(3) $$
\begin{aligned}
E(\boldsymbol{M}_{k+1}|\mathcal{F}_k) &= E(A_1^{-1} A_2^{-1} \cdots A_{k+1}^{-1} \boldsymbol{\xi}_{k+1}|\mathcal{F}_k) \\
&= A_1^{-1} A_2^{-1} \cdots A_{k+1}^{-1} E(\boldsymbol{\xi}_{k+1}|\mathcal{F}_k) \\
&= A_1^{-1} A_2^{-1} \cdots A_{k+1}^{-1} A_{k+1} \boldsymbol{\xi}_k = A_1^{-1} A_2^{-1} \cdots A_k^{-1} \boldsymbol{\xi}_k = \boldsymbol{M}_k.
\end{aligned}
$$

Now, to extract the benefit from the martingale $\{\boldsymbol{M}_k\}$, one just needs to make it more explicit.

Here one is fortunate; an easy induction confirms that $A_1^{-1}A_2^{-1}\cdots A_k^{-1}$ is given by the upper-triangular matrix:

$$\frac{1}{n-k}\begin{pmatrix} \dfrac{n(n-1)}{n-k-1} & -\dfrac{2knM}{n-k-1} & \dfrac{kn}{n-k-1} & \dfrac{k(k+1)M^2-knB}{n-k-1} \\[2mm] 0 & n & 0 & -kM \\[2mm] 0 & 0 & n & -kB \\[2mm] 0 & 0 & 0 & n-k \end{pmatrix}.$$

Now, the coordinates of the vector martingale $\boldsymbol{M}_k = (M_{1,k}, M_{2,k}, M_{3,k}, M_{4,k})^\top$ are martingales in their own right, and it is worthwhile to examine them individually.

The fourth coordinate just gives the trivial martingale $M_{4,k} \equiv 1$, but the other coordinates are much more interesting. The second and third coordinates give us two useful — but known — martingales,

(4) $\qquad M_{2,k} = (nS_k - kM)/(n-k) \quad$ and $\quad M_{3,k} = (nT_k - kB)/(n-k).$

Actually, we have only one martingale here; one gets the martingale $\{M_{3,k}\}$ from the martingale $\{M_{2,k}\}$ if one replaces $\mathcal{X}$ with the set of squares $\mathcal{X}' = \{x_1^2, x_2^2, ..., x_n^2\}$. The martingale $\{M_{2,k}\}$ is given in Serfling (1974) and Stout (1974, p. 147). The earliest source we could identify for the martingale $\{M_{2,k}\}$ is Garsia (1968, p. 82).

To get a martingale that is more novel (and less transparent), we only need to consider the first coordinate of the vector martingale $\boldsymbol{M}_k$. One can write this coordinate explicitly as

$$M_{1,k} = \frac{n(n-1)}{(n-k)(n-k-1)}S_k^2 - \frac{2knM}{(n-k)(n-k-1)}S_k$$
$$+ \frac{kn}{(n-k)(n-k-1)}T_k + \frac{k(k+1)M^2-knB}{(n-k)(n-k-1)}.$$

This may seem complicated at first sight, but there is room for simplification. In many situations it is natural to assume that $M = S_n = 0$, and, in that case, $M_{1,k}$ reduces to the more manageable martingale sequence that we denote by

(5) $\quad \widetilde{M}_k = \dfrac{1}{(n-k)(n-k-1)}\left[(n-1)S_k^2 - (B-T_k)k\right], \qquad 1 \le k \le n-2.$

We will give several applications of this martingale to permutation inequalities, and, in particular, we use it in Section 6 to get a new bi-quadratic maximal inequality for permuted arrangements. We should also note that this martingale also has a potentially useful monotonicity property. Specifically, T_k is monotone increasing, so surgery on T_k (say replacing T_k by $k^\alpha T_k$) yields a rich family of submartingales or supermartingales.

3. Second Example of the Construction

As before, we assume that $X_1, X_2, ..., X_n$ is a sequential sample taken without replacement from $\mathcal{X}$, and we impose the centering condition $S_n = M = 0$. Further, we consider a system of "multipliers" a_k, $1 \le k \le n$ where each a_k is assumed to be an $\mathcal{F}_{k-1}$ measurable random variable. Since our measure space is finite, these non-anticipating random variables are automatically bounded.

The basic building block for our next collection of martingales is the sequence of random variables defined by setting

$$W_k = a_1 X_1 + a_2 X_2 + \cdots + a_k X_k \quad \text{for } 1 \leq k \leq n.$$

The immediate task is to find a martingale that has W_k (or a deterministic multiple of W_k) as a summand.

As before, we begin by calculating the one-step conditional expectations:

$$E[W_{k+1}|\mathcal{F}_k] = W_k - \frac{a_{k+1}}{n-k} S_k \quad \text{and} \quad E[S_{k+1}|\mathcal{F}_k] = \frac{n-k-1}{n-k} S_k.$$

If we introduce the vector $\boldsymbol{\eta}_k = (W_k, S_k)^\top$ we have

$$E[\boldsymbol{\eta}_{k+1}|\mathcal{F}_k] = A_{k+1}\boldsymbol{\eta}_k \quad \text{where } A_{k+1} = \begin{pmatrix} 1 & -\dfrac{a_{k+1}}{n-k} \\ 0 & \dfrac{n-k-1}{n-k} \end{pmatrix}.$$

Inversion is now especially easy, and we note that for

$$A_k = \begin{pmatrix} 1 & -\dfrac{a_k}{n-k+1} \\ 0 & \dfrac{n-k}{n-k+1} \end{pmatrix} \quad \text{we have} \quad A_k^{-1} = \begin{pmatrix} 1 & \dfrac{a_k}{n-k} \\ 0 & \dfrac{n-k+1}{n-k} \end{pmatrix},$$

so induction again confirms the critical inverse:

$$A_1^{-1} A_2^{-1} \cdots A_k^{-1} = \begin{pmatrix} 1 & \dfrac{a_1 + a_2 + \cdots + a_k}{n-k} \\ 0 & \dfrac{n}{n-k} \end{pmatrix}.$$

The general recipe (3) then gives us a new martingale:

$$(6) \qquad M_k = W_k + \frac{a_1 + a_2 + \cdots + a_k}{n-k} S_k \qquad 1 \leq k < n$$

If one were somehow to consider the expression for $\{M_k : 1 \leq k < n\}$ *ex cathedra*, one could subsequently verify the martingale property by a direct calculation. Thus, in this instance, linear algebra has served us mainly as a tool for discovery.

Some of the benefits of the martingale (6) are brought to life through interesting choices for the non-anticipating factors $\{a_k : 1 \leq k \leq n\}$. For, example, if we take $a_1 = 0$ and set $a_k = X_{k-1}$ for $k \geq 2$, we find a curious quadratic martingale

$$M_k = X_1 X_2 + X_2 X_3 + \cdots + X_{k-1} X_k + \frac{(X_1 + \cdots + X_{k-1})(X_1 + \cdots + X_k)}{n-k}.$$

One is unlikely to have hit upon this martingale without the benefit of a systematic construction. For the moment, no application of this martingale comes to mind, but it does seem useful to know that there is such a simple quadratic martingale. Perhaps a nice application is not too far away.

Out main purpose here is to expose the general construction (3), but, through the illustrations just given, we now have several martingales that speak directly to permutation inequalities — an extensive subject where martingales have traditionally

been part of the toolkit. As we explore what can be done with the new martingales (5) and (6), we will also address some of the classic results on permutation inequalities.

4. A Permutation Inequality

In the first application, we quickly check what can be done with the martingale $M_{2,k} = (nS_k - kM)/(n - k)$ given by (4) in our first construction. To keep the formulas simple, we impose a standing assumption,

$$(7) \qquad\qquad S_n = M = 0,$$

so, $S_k/(n - k)$ is an $\{\mathcal{F}_k\}$-martingale with expectation zero. The Doob-Kolmogorov L^p maximal inequality (see e.g. Shiryaev (1995, p. 493)) tells one that for any L^p, $1 < p < \infty$, martingale $\{M_k : 1 \leq k \leq n\}$ one has the bound

$$|| \max_{1 \leq k \leq n} |M_k| ||_p \leq \frac{p}{1 - p} ||M_n||_p,$$

so, taking $p = 2$, $M_k = S_k/(n - k)$, and $1 \leq k \leq n - 1$, we have

$$E \max_{1 \leq k \leq n-1} \left| (n - k)^{-1} S_k \right|^2 \leq 4ES_{n-1}^2.$$

This can be simplified by noting that

$$ES_{n-1}^2 = E(M - X_n)^2 = EX_n^2 = EX_1^2 = B/n,$$

so, in the end, we have

$$(8) \qquad\qquad E \max_{1 \leq k \leq n-1} \left| (n - k)^{-1} S_k \right|^2 \leq 4B/n.$$

One could immediately transcribe this as a permutation inequality, but first we put it in a form that seems more natural for applications.

For any fixed permutation σ, the distribution of the vector $(X_1, X_2, \ldots, X_n)$ is the same as distribution the vector $(X_{\sigma(1)}, X_{\sigma(2)}, \ldots, X_{\sigma(n)})$ (cf. Feller (1971, p. 228)), so, in particular the distribution of $(X_1, X_2, \ldots, X_n)$ is the same as the distribution of $(X_n, X_{n-1}, \ldots, X_1)$. By the centering assumption (7), we know

$$(X_1 + X_2 + \cdots + X_k)^2 = (X_{k+1} + X_{k+2} + \cdots + X_n)^2,$$

so applying both observations gives us the identity

$$E \max_{1 \leq k \leq n-1} \left| (n - k)^{-1} S_k \right|^2 = E \max_{1 \leq k \leq n-1} |S_k/k|^2 .$$

Using this bound in (8) then gives us

$$(9) \qquad\qquad E \max_{1 \leq k \leq n-1} |S_k/k|^2 \leq 4B/n,$$

so by the sampling model and the definition of B, we come to an attractive inequality for the maximum of averages drawn sequentially from a randomly permutated sample.

Proposition 1 (Max-Averages Inequality)*. For real numbers $\{x_1, x_2, ..., x_n\}$ with $x_1 + x_2 + \cdots + x_n = 0$, we have*

$$(10) \qquad\qquad \frac{1}{n!} \sum_{\sigma} \max_{1 \leq k \leq n} \left\{ \frac{1}{k} \sum_{i=1}^{k} x_{\sigma(i)} \right\}^2 \leq \frac{4}{n} \sum_{i=1}^{n} x_i^2.$$

The traditional version of Hardy's inequality (cf. Hardy (1920) and Steele (2004), pp. 169) tell us that for real $a_1, a_2, \ldots, a_n$ we have

$$(11) \qquad \sum_{k=1}^{n} \left\{ \frac{1}{k} \sum_{i=1}^{k} a_i \right\}^2 \le 4 \sum_{i=1}^{n} a_i^2,$$

and by applying this to $a_1 = x_{\sigma(1)}$, $a_2 = x_{\sigma(2)}$, ..., $a_n = x_{\sigma(n)}$ for each permutation σ, one has

$$(12) \qquad \max_{\sigma} \sum_{k=1}^{n} \left\{ \frac{1}{k} \sum_{i=1}^{k} x_{\sigma(i)} \right\}^2 \le 4 \sum_{i=1}^{n} x_i^2.$$

It is well known (and easy to prove) that the constant in Hardy's inequality is best possible, so by the analogy between (12) and (10) we conjecture that the constant in (10) is best possible. Finally, we note that in Section 8 we give two applications of the Max-Averages Inequality (10).

5. Exchangeability and a Folding Device

One also has the possibility of using exchangeability more forcefully; in particular, one can exploit exchangeability around the center of the sample. For a preliminary illustration of this possibility, we fix $1 \le m < n$ and note that the martingale property (5) of $\widetilde{M}_k$ gives us

$$ES_m^2 = E \frac{m}{n-1}[B - T_m] = \frac{m}{n-1}[B - ET_m].$$

The martingale property (4) of $M_{3,k}$ also gives us $ET_m = (m/n)B$, so we have

$$(13) \qquad ES_m^2 = \frac{m(n-m)}{n(n-1)} B,$$

a fact that one can also get by bare hands, though perhaps not so transparently.

We can also use the martingale $M_{2,k} = (nS_k - kM)/(n-k) = nS_k/(n-k)$ here. Since we have $M = 0$, the L^2 maximal inequality for this martingale and the identity (13) give us

$$(14) \qquad E \max_{1 \le k \le m} \left| \frac{S_k}{n-k} \right|^2 \le \frac{4ES_m^2}{(n-m)^2} = \frac{4mB}{n(n-1)(n-m)}.$$

Since $(n-k)^2 \le n^{-1}(n-1)$ for $1 \le k \le m$, the bound (14) implies the weaker, but simpler, bound

$$(15) \qquad E \max_{1 \le k \le m} S_k^2 \le \frac{4mB}{n-m} \qquad \text{for } 1 \le m < n.$$

The idea now is to use symmetry to exploit the fact that (15) holds for many choices of m. To begin, we note that we always have the crude bound

$$(16) \qquad E \max_{1 \le k \le n} S_k^2 \le E \max_{1 \le k \le m} S_k^2 + E \max_{m < k \le n} S_k^2.$$

Typically this is useless, but here it points to a useful observation; we can use the same bounds on the second terms that we used on the first. This is the "folding device" of the section heading.

Specifically, since $M = 0$ we have

$$|S_k| = |S_n - S_k| = |X_n + X_{n-1} + \cdots + X_{k+1}|,$$

so, by exchangeability, we see that the second summand of (16) is bounded by $4(n-m)B/m$. Thus, we have for all $1 \leq m < n$ that

$$(17) \qquad E \max_{1 \leq k \leq n} S_k^2 \leq E \max_{1 \leq k \leq m} S_k^2 + E \max_{m < k \leq n} S_k^2 \leq 4B \left(\frac{m}{n-m} + \frac{n-m}{m} \right).$$

Here, if we just consider $n \geq 8$ and choose m as close as possible to $n/2$, then we see that the worst case occurs when $n = 9$ and $m = 4$; so we have the bound

$$(18) \qquad E \max_{1 \leq k \leq n} S_k^2 \leq E \max_{1 \leq k \leq m} S_k^2 + E \max_{m < k \leq n} S_k^2 \leq (41/5)B \quad \text{for } n \geq 8.$$

On the other hand, Cauchy's inequality gives us $S_k^2 \leq kB \leq nB$, so the bound (18) is trivially true for $n \leq 7$. Combining these ranges gives us a centralized version of an inequality originating with A. Garsia (cf. Stout (1974, pp. 145–148)).

Proposition 2 (A. Garsia). *For any set of real numbers $\{x_1, x_2, ..., x_n\}$ with sum $x_1 + x_2 + \cdots + x_n = 0$ we have*

$$(19) \qquad \frac{1}{n!} \sum_\sigma \max_{1 \leq k \leq n} | \sum_{i=1}^{k} x_{\sigma(i)} |^2 \leq (8 + \frac{1}{5}) \sum_{i=1}^{n} x_i^2,$$

where the sum is taken over all possible permutations of $\{x_1, x_2, ..., x_n\}$.

Here we have paid some attention to the constant in this inequality, but already from the crude bound (16) one knows that the present approach is not suited to the derivation of a best possible result. Our intention has been simply to illustrate what one can do with reasonable care and robust approach that uses the tools provided by our general martingale construction.

Currently the best value for the constant in the Garsia inequality (19) is due to Chobanyan (1994, Corollary 3.3) where the stunning value of 2 is obtained. Moreover, Chobanyan and Salehi (2001, Corollary 2.8)) have a much more general inequality which also gives a constant of 2 when specialized to our situation.

Here we should also note that in all inequalities of this type have both a centralized version where $M = 0$ and non-centralized version where M is unconstrained. One can pass easily between the versions (see e.g. Stout (1974, p. 147) or Garsia (1970, p. 93)). The centralized versions are inevitably simpler to state, so we have omitted discussions of the non-centralized versions.

6. QUADRATIC PERMUTATION INEQUALITY

To apply the L^2 maximal inequality to the martingale $\{\widetilde{M}_k : 1 \leq k \leq n-2\}$ that was discovered by our first construction (5), one needs an estimate of the bounding term $4E[\widetilde{M}_{n-2}^2]$ in the Doob-Kolmogorov inequality. There are classical formulas for the moments for sampling without replacement that simplify this task. These formulas were known to Isserlis (1931), if not before, but they are perhaps easiest to derive on one's own.

Lemma 3. *If X_1, X_2, X_3, X_4 are draw without replacement from $\mathcal{X} = \{x_1, x_2, ..., x_n\}$ where $x_1 + x_2 + \cdots + x_n = 0$, then we have the moments*

$$E(X_1 X_2 X_3 X_4) = \frac{3B^2 - 6Q}{n(n-1)(n-2)(n-3)}, \quad E(X_1^2 X_2 X_3) = \frac{2Q - B^2}{n(n-1)(n-2)},$$

$$E(X_1^2 X_2^2) = \frac{B^2 - Q}{n(n-1)}, \quad E(X_1^3 X_2) = -\frac{Q}{n(n-1)}, \quad and \quad E(X_1^4) = Q/n,$$

where, as usual, we set $B = x_1^2 + x_2^2 + \cdots + x_n^2$ *and* $Q = x_1^4 + x_2^4 + \cdots + x_n^4$.

Now, to calculate $4E[\widetilde{M}_{n-2}^2]$, we first note that

$$S_{n-1} = -X_n \quad \text{and} \quad S_{n-2} = -X_{n-1} - X_n,$$

so just expanding the definition of $\widetilde{M}_{n-2}$ gives us

$$
\begin{aligned}
4E[\widetilde{M}_{n-2}^2] =& E\left[\{(n-1)S_{n-2}^2 - (n-2)(B - T_{n-2})\}^2\right] \\
=& E\left[\{(n-1)(X_{n-1} + X_n)^2 - (n-2)(X_{n-1}^2 + X_n^2)\}^2\right] \\
=& E\left[X_{n-1}^4 + X_n^4 + (4n^2 - 8n + 6)X_{n-1}^2 X_n^2 \right. \\
& \left. + 4(n-1)X_{n-1}^3 X_n + 4(n-1)X_{n-1}X_n^3\right].
\end{aligned}
$$

By Lemma 3 and exchangeability, one then finds after some algebra that the L^2 maximal inequality takes the form

$$(20) \qquad E \max_{1 \le k \le n-2} \widetilde{M}_k^2 \le 4E[\widetilde{M}_{n-2}^2] = \frac{4n^2 - 8n + 6}{n(n-1)}B^2 - \frac{4n^2 - 2n}{n(n-1)}Q.$$

The only task left is to reframe this inequality so that it easy to apply as a permutation inequality. Here it is useful to observe that

$$(21) \qquad \frac{4n^2 - 8n + 6}{n(n-1)} < 4 \quad \text{and} \quad \frac{4n^2 - 2n}{n(n-1)} > 4 \qquad \text{for } n \ge 2,$$

moreover, these are not wasteful bounds; they are essentially sharp for large n. When we apply these bounds in (20) we get the nicer bound,

$$(22) \qquad E \max_{1 \le k \le n-2} \left|\frac{(n-1)S_k^2 - k(B - T_k)}{(n-k)(n-k-1)}\right|^2 \le 4[B^2 - Q].$$

This brings us closer to our goal, but further simplification is possible if we note that by reverse sampling the left side of (22) can also be written as

$$
\begin{aligned}
E \max_{1 \le k \le n-2} \left|\frac{(n-1)S_k^2 - k(B - T_k)}{(n-k)(n-k-1)}\right|^2 =& E \max_{2 \le k \le n-1} \left|\frac{(n-1)S_k^2 - (n-k)T_k}{k(k-1)}\right|^2 \\
=& E \max_{2 \le k \le n} \left|\frac{(n-1)S_k^2 - (n-k)T_k}{k(k-1)}\right|^2.
\end{aligned}
$$

In permutation terms, the bound (22) establishes the following proposition.

Proposition 4 (Quadratic Permutation Inequality). *For any set of real numbers* $\{x_1, x_2, ..., x_n\}$ *with* $x_1 + x_2 + \cdots + x_n = 0$ *we have*

$$\frac{1}{n!}\sum_{\sigma} \max_{2 \le k \le n} \left|\frac{\left(\sum_{i=1}^k x_{\sigma(i)}\right)^2 - \frac{n-k}{n-1}\sum_{i=1}^k x_{\sigma(i)}^2}{k(k-1)}\right|^2 \le \frac{4}{(n-1)^2}\left\{\left(\sum_{i=1}^n x_i^2\right)^2 - \sum_{i=1}^n x_i^4\right\},$$

where the first sum is taken over all possible permutations of $\{x_1, x_2, ..., x_n\}$.

At first glance, this may seem complicated, but the components are all readily interpretable and the inequality is no more complicated than it has to be. The main observation is that our general construction (2) brought us here in a completely straightforward way. Without such an on-ramp, one is unlikely to imagine any

bound of this kind — and *relative* simplicity. Moreover, the inequality does have intuitive content, and this content is made more explicit in the next section.

7. The Discrete Bridge and Further Folding

Here we set $x_i = 1$ for $1 \leq i \leq m$ and take $x_i = -1$ for $m + 1 \leq i \leq 2m$. We then let X_i, $1 \leq i \leq 2m$, denote samples that are drawn without replacement from the set $\mathcal{X} = \{x_1, x_2, \ldots, x_{2m}\}$. If we put $S_0 = 0$ and denote the usual partial sums by S_k, $1 \leq k \leq 2m$, then $S_{2m} = 0$ and the process $\{S_k : 0 \leq k \leq 2m\}$ is a discrete analog of the Brownian bridge. Alternatively, one can view this process as simple random walk that is conditioned to return to 0 at time $2m$.

For $\{\widetilde{M}_k\}$, the martingale from the first construction (5), we now have $T_k \equiv k$, so we have the simple representation

$$(23) \qquad \widetilde{M}_k = \frac{(2m-1)S_k^2 - (2m-k)k}{(2m-k)(2m-k-1)}, \qquad 1 \leq k \leq 2m - 2.$$

By Lemma 3 with $B = Q = 2m$ and by (13), we also find that at the (left) mid-point m of our process we have the nice relations

$$(24) \qquad E[S_m^2] = m^2/(2m-1) \quad \text{and} \quad E[S_m^4] = \frac{3m^4 - 4m^3}{4m^2 - 8m + 3}.$$

These give us a rational formula for $E[\widetilde{M}_m^2]$, but for the moment, we just use the partial simplification

$$E[\widetilde{M}_m^2] = E\left| \frac{(2m-1)S_m^2 - (2m-m)m}{(2m-m)(2m-m-1)} \right|^2$$

$$= \frac{1}{m^2(m-1)^2}\left[(2m-1)^2 ES_m^4 - m^4 \right].$$

For $1 \leq k \leq m$ we have the trivial bound

$$1 \leq \max_{1 \leq k \leq m} \left| \frac{(2m)(2(m-1))}{(2m-k)(2m-k-1)} \right|^2,$$

so the L^2 maximal inequality applied to the martingale (23) gives us

$$(25) \qquad E \max_{1 \leq k \leq m} \left| (2m-1)S_k^2 - (2m-k)k \right|^2 \leq 64\left[(2m-1)^2 ES_m^4 - m^4 \right].$$

We can now take advantage of a symmetry that is special to the discrete bridge. Since $S_{2m} = 0$ we have $(X_1 + \cdots + X_k)^2 = (X_{2m} + \cdots + X_{k+1})^2$, and by exchangeability the vectors $(X_1, \ldots, X_m)$ and $(X_{2m}, \ldots, X_{m+1})$ have the same distribution. Also, the value $(2m-k)k$ is "invariant" in the following sense: if we substitute for k (the number of summands in $X_1 + X_2 + \cdots + X_k$) the value $2m - k$ (the number of summands in $X_{2m} + X_{2m-1} + \ldots + X_{k+1}$), then the symmetric quantity $(2m-k)k$ is unchanged. As a consequence, the random variables

$$\max_{1 \leq k \leq m} \left| (2m-1)S_k^2 - (2m-k)k \right|^2 \quad \text{and} \quad \max_{m \leq k \leq 2m-1} \left| (2m-1)S_k^2 - (2m-k)k \right|^2$$

are equal in distribution.

We can then apply (25) twice to obtain

$$E \max_{1 \leq k \leq 2m-1} \left| (2m-1)S_k^2 - (2m-k)k \right|^2 \leq 128\left[(2m-1)^2 ES_m^4 - m^4 \right].$$

From this bound and the formula (24) for ES_m^4, we then have

$$E \max_{1 \leq k \leq 2m-1} \left| S_k^2 - k\frac{2m-k}{2m-1} \right|^2 \leq 128 \frac{4m^3 - 8m^2 + 4m}{8m^3 - 20m^2 + 14m - 3} m^2,$$

and, for any $m \geq 2$ we have $(4m^3 - 8m^2 + 4m)/(8m^3 - 20m^2 + 14m - 3) < 1$. In the end, we have the following proposition.

Proposition 5 (Quadratic Permutation Inequality for Discrete Bridge)**.** *For the set* $\mathcal{X} = \{x_1, x_2, \ldots, x_{2m}\}$ *with* $x_i = 1$ *for* $1 \leq i \leq m$ *and* $x_i = -1$ *for* $m + 1 \leq i \leq 2m$ *one has*

$$(26) \qquad \frac{1}{(2m)!} \sum_\sigma \max_{1 \leq k \leq 2m-1} \left| \left(\sum_{i=1}^k x_{\sigma(i)} \right)^2 - k\frac{2m-k}{2m-1} \right|^2 \leq 128m^2,$$

where the first sum is taken over all possible permutations.

For the terms of the squared discrete bridge process $\{S_k^2 : 0 \leq i \leq 2m\}$ have the expectations $ES_k^2 = k(2m - k)/(2m - 1)$, so (26) gives us a rigorous bound on the maximum deviation of the square S_k^2 of a discrete bridge from its expected value ES_k^2. Here, the order $O(m^2)$ of the bound cannot be improved, but, if necessity called, one may be able to improve on ungainly constant 128.

8. Permuted Sums with Fixed Weights

In Section 3, we considered the weighted sums

$$W_k = a_1 X_1 + a_2 X_2 + \cdots + a_k X_k \quad \text{for } 1 \leq k \leq n.$$

and we found that the process

$$(27) \qquad M_k = W_k + \frac{a_1 + a_2 + \cdots + a_k}{n - k} S_k \qquad 1 \leq k < n$$

is a martingale whenever the multipliers are non-anticipating random variables (i.e. whenever a_k is $\mathcal{F}_{k-1}$-measurable for each $1 \leq k \leq n$). Here we will show that this martingale has informative uses even when one simply takes the multipliers to be fixed real numbers.

First we introduce some shorthand. For $1 \leq k \leq n$ we will write

$$(28) \qquad \alpha_1(k) = a_1 + a_2 + \cdots + a_k \quad \text{and} \quad \alpha_2(k) = a_1^2 + a_2^2 + \cdots + a_k^k.$$

Also, we will be most interested in multiplier vectors $a = (a_1, a_2, \ldots, a_n)$ for which we have some control of the quantity

$$(29) \qquad V_n(a) \stackrel{\text{def}}{=} \max_{1 \leq k \leq n-1} \alpha_1^2(k)/\alpha_2(n),$$

which is a measure of cancelation among the multipliers. A leading example worth keeping in mind is the sequence $a_k = (-1)^{k+1}$, $1 \leq k \leq n$, for which we have $V_n(a) = 1/n$. We will revisit the measure $V_n(a)$ after we derive a moment bound.

Lemma 6. *For the martingale* $\{M_k\}$ *defined by equation* (27)*, we have*

$$(30) \qquad E[M_k^2] = \frac{1}{n-1}\alpha_2(k)B + \frac{1}{(n-1)(n-k)}\alpha_1^2(k)B.$$

Proof. Simply squaring (27), we have

$$|M_k|^2 = W_k^2 + \frac{\alpha_1^2(k)}{(n-k)^2} S_k^2 + 2\frac{\alpha_1(k)}{n-k} W_k S_k,$$

so we just need to find EW_k^2, ES_k^2 and $EW_k S_k$. We know ES_k^2 from (13), and we have $EX_i^2 = B/n$ and $EX_i X_j = -B/(n(n-1))$ so

$$EW_k^2 = E\Big[\sum_{i=1}^{k} a_i^2 X_i^2 + \sum_{1 \le i,j \le k, i \ne j} a_i a_j X_i X_j\Big]$$

$$= \alpha_2(k)\frac{B}{n} + \big[\alpha_1^2(k) - \alpha_2(k)\big]\left[-\frac{B}{n(n-1)}\right] = \alpha_2(k)\frac{B}{n-1} - \alpha_1^2(k)\left[\frac{B}{n(n-1)}\right].$$

Similarly, we have

$$EW_k S_k = E\Big[\sum_{i=1}^{k} a_i X_i^2 + \sum_{i=1}^{k} a_i \sum_{1 \le j \le k, j \ne i} X_i X_j\Big]$$

$$= \alpha_1(k)\frac{B}{n} - (k-1)\alpha_1(k)\left[\frac{B}{n(n-1)}\right] = \alpha_1(k)\frac{(n-k)B}{n(n-1)},$$

so summing up the terms completes the proof of the lemma. □

From this lemma, the L^2 maximal inequality gives us

$$(31) \qquad E \max_{1 \le k \le n-1} M_k^2 \le 4E[M_{n-1}]^2 = 4[\alpha_2(n-1) + \alpha_1^2(n-1)]\frac{B}{n-1},$$

but to extract the full benefit from the bound (31) we need to relate it to the weighted sum W_k. Using (27) we can write W_k as the difference between M_k and $\alpha_1(k)S_k/(n-k)$, so from the trivial bound $(x+y)^2 \le 2x^2 + 2y^2$ we have

$$\max_{1 \le k \le n-1} |W_k|^2 \le 2 \max_{1 \le k \le n-1} |M_k|^2 + 2 \max_{1 \le k \le n-1} \alpha_1^2(k)\left|\frac{S_k}{n-k}\right|^2$$

$$\le 2 \max_{1 \le k \le n-1} |M_k|^2 + 2 \max_{1 \le k \le n-1} \alpha_1^2(k) \cdot \max_{1 \le k \le n-1} \left|\frac{S_k}{n-k}\right|^2.$$

The second step may look wasteful, but now we can apply both (31) and the Max-Averages inequality (8) from Section 4 to obtain

$$E \max_{1 \le k \le n-1} |W_k|^2 \le 8\Big[\alpha_2(n-1) + \alpha_1^2(n-1)\Big]\frac{B}{n-1} + 8 \max_{1 \le k \le n-1} \alpha_1^2(k)\frac{B}{n}.$$

Note that we also have

$$\max_{1 \le k \le n} |W_k|^2 = \max\left(\max_{1 \le k \le n-1} |W_k|^2, (W_{n-1} + a_n X_n)^2\right)$$

$$\le \max\left(\max_{1 \le k \le n-1} |W_k|^2, 2W_{n-1}^2 + 2a_n^2 X_n^2\right)$$

$$\le 2 \max_{1 \le k \le n-1} |W_k|^2 + 2a_n^2 X_n^2,$$

and the bottom line is that

$$(32) \qquad E\left\{\max_{1 \le k \le n} |W_k|^2\right\} = \frac{1}{n!}\sum_{\sigma} \max_{1 \le k \le n} \left|\sum_{i=1}^{k} a_i x_{\sigma(i)}\right|^2$$

is bounded by the lengthy (but perfectly tractable) sum

$$(33) \qquad 16\Big[\alpha_2(n-1) + \alpha_1^2(n-1)\Big]\frac{B}{n-1} + 16\max_{1\le k\le n-1}\alpha_1^2(k)\frac{B}{n} + 2a_n^2\frac{B}{n}.$$

To make this concrete, note that for $a_k = (-1)^{k+1}$ the bound on (32) that we get from (33) is simply $(16n/(n-1) + 18/n)B$. The ratio $16n/(n-1) + 18/n$ decreases to 16, and for $n = 18$ the upper bound is equal to $(17 + 16/17)B$. By Cauchy's inequality, we have $|W_k|^2 \le kB$ for all k, so the bound $(17 + 16/17)B$ also holds for all $n \le 17$. When we assemble the pieces, we have a permutation maximal inequality for sums with alternating signs.

Proposition 7. *For real numbers $\{x_1, x_2, ..., x_n\}$ with $x_1 + x_2 + \cdots + x_n = 0$ we have*

$$(34) \qquad \frac{1}{n!}\sum_\sigma \max_{1\le k\le n} |\sum_{i=1}^k (-1)^i x_{\sigma(i)}|^2 \le \left(17 + \frac{16}{17}\right)\sum_{i=1}^n x_i^2$$

where the sum is taken over all permutations of $\{x_1, x_2, ..., x_n\}$.

The argument that leads to (34) is useful for more than just alternating sums; it is effective whenever there is meaningful cancelation in $a = (a_1, a_2, \ldots, a_n)$. Specifically, the bound (33) is always dominated by

$$16\alpha_2(n) + 32\max_{1\le k\le n-1}\alpha_1^2(k),$$

so our proof also gives us a more general bound.

Proposition 8. *For sets of real numbers $\{a_1, a_2, ..., a_n\}$ and $\{x_1, x_2, ..., x_n\}$ such that $x_1 + x_2 + \cdots + x_n = 0$ we have*

$$(35) \qquad \frac{1}{n!}\sum_\sigma \max_{1\le k\le n} |\sum_{i=1}^k a_i x_{\sigma(i)}|^2 \le \frac{16}{n-1}\{1 + 2V_n(a)\}\sum_{i=1}^n a_i^2 \sum_{i=1}^n x_i^2,$$

where the sum is taken over all permutations of $\{x_1, x_2, ..., x_n\}$ and where $V_n(a)$ defined in (29).

This inequality shows that there are concrete benefits to introducing $V(a)$. For example, by a sustained and subtle argument, Garsia (1970, 3.7.20, p. 92) arrived at a version of our inequality where the coefficient $16\{1 + 2V_n(a)\}$ is replaced by 80. Now, for uniform multipliers, one has $V_n(a) = O(n)$ and Garsia's inequality is greatly superior to our bound (35). On the other hand, for multipliers that satisfy the cancelation property $V_n(a) \to 0$, the present bound eventually meets and beats Garsia's bound. In particular, for multipliers given by alternating signs, the constant of (35) is just $17 + 16/17$.

9. Weighted Sums and Folding

By many accounts, the permutation maximal inequality of Garsia (1970, p. 86) is the salient result in the theory of permutation inequalities, so it is a natural challenge to see if it can be proved by the robust martingale methods that follow from our martingale constructions. We give a proof of this kind — without appeal to $V_n(a)$. Once the proof is complete, we address the differences between Garsia's inequality and the present bound with its curious constant.

Proposition 9. *For real numbers $\{a_1, a_2, ..., a_n\}$ and real numbers $\{x_1, x_2, ..., x_n\}$ such that $x_1 + x_2 + \cdots + x_n = 0$ we have*

$$(36) \qquad \frac{1}{n!} \sum_\sigma \max_{1 \le k \le n} \left| \sum_{i=1}^k a_i x_{\sigma(i)} \right|^2 \le \left(80 + \frac{4}{205}\right) \sum_{i=1}^n a_i^2 \sum_{i=1}^n x_i^2 / (n-1),$$

where the sum is taken over all possible permutations of $\{x_1, x_2, ..., x_n\}$.

Proof. First, for all $1 \le k \le n$, Cauchy's inequality gives us $\alpha_1^2(k) \le k\alpha_2(k)$. Trivially one has $\alpha_2(k) \le \alpha_2(n)$, so by Lemma 6 we have the bound

$$(37) \qquad E[M_k]^2 \le \left(1 + \frac{k}{n-k}\right) \frac{\alpha_2(n)B}{n-1} \qquad \text{for } 1 \le k \le n-1.$$

Now, just from the definition (27) of M_k, we can write W_k as a difference between M_k and $\alpha_1(k)S_k/(n-k)$. We can then use the crude bound $(x+y)^2 \le 2x^2 + 2y^2$ and Cauchy's inequality to get

$$\max_{1 \le k \le m} |W_k|^2 \le 2 \max_{1 \le k \le m} |M_k|^2 + \alpha_1^2(k) \left| \frac{S_k}{n-k} \right|^2$$

$$\le 2 \max_{1 \le k \le m} |M_k|^2 + 2 \max_{1 \le k \le m} \alpha_1^2(k) \left| \frac{S_k}{n-k} \right|^2$$

$$\le 2 \max_{1 \le k \le m} |M_k|^2 + (2m)\alpha_2(n) \max_{1 \le k \le m} \left| \frac{S_k}{n-k} \right|^2.$$

Long ago, in (13), we calculated ES_m^2, so here we can apply the L^2 maximal inequality to both martingales $\{S_k/(n-k)\}$ and $\{M_k\}$ to get the bound

$$(38) \qquad E \max_{1 \le k \le m} W_k^2 \le 8 \left(1 + \frac{m}{n-m} + \frac{m^2}{n(n-m)}\right) \frac{\alpha_2(n)B}{n-1}.$$

Such an inequality for $1 \le m < n$ suggests the possibility of folding. To pursue this we first note

$$\max_{1 \le k \le n} W_k^2 = \max \left[\max_{1 \le k \le m} W_k^2, \ \max_{m < k \le n} W_k^2 \right]$$

$$= \max \left[\max_{1 \le k \le m} W_k^2, \ \max_{m < k \le n} |W_m + a_{m+1}X_{m+1} + \cdots + a_k X_k|^2 \right]$$

$$\le \max \left[\max_{1 \le k \le m} W_k^2, \ 2W_m^2 + 2 \max_{m < k \le n} |a_{m+1}X_{m+1} + \cdots + a_k X_k|^2 \right]$$

$$(39) \qquad \le 2 \left[\max_{1 \le k \le m} W_k^2 + \max_{m < k \le n} |a_{m+1}X_{m+1} + \cdots + a_k X_k|^2 \right].$$

By exchangeability, one expects the second maximum has a bound like the one derived in (38). To make this explicit, one simply needs to replace m by $n-m$ in the upper bound of (38). Doing so gives us the sister bound

$$E \max_{m < k \le n} |a_{m+1}X_{m+1} + \cdots + a_k X_k|^2 \le 8 \left(1 + \frac{(n-m)}{m} + \frac{(n-m)^2}{nm}\right) \frac{\alpha_2(n)B}{n-1}.$$

By our bounds on the two addends of (39), we then have

$$E \max_{1 \le k \le n} W_k^2 \le 16 \left(2 + \frac{m}{n-m} + \frac{n-m}{m} + \frac{m^2}{n(n-m)} + \frac{(n-m)^2}{nm}\right) \frac{\alpha_2(n)B}{n-1}.$$

It only remains to take $m = \lfloor n/2 \rfloor$ and to attend honestly to the consequences. If n is even, the constant that multiplies $\alpha_2(n)B/(n-1)$ is exactly 80. For odd n, the constant approaches 80 from above as n increases to infinity, and for $n = 81$ the constant is $80 + 4/205$. Furthermore, Cauchy's inequality gives us

$$W_k^2 \le \alpha_2(k)B \le \alpha_2(n)B \quad \text{for all } 1 \le k \le n,$$

and we have $\alpha_2(n)B \le (80 + 4/205)\alpha_2(n)B/(n-1)$ for all $n \le 81$. So, in the end, we come to (36), our permutation maximal inequality with general weights. $\square$

We would greet this result with some fanfare except that Garsia (1970, 3.7.20, p. 92) gives this bound with the constant 80. Still, one needs to keep in mind that we have pursued this derivation only to illustrate the usefulness of the martingales that are given by our general linear algebraic construction. Perhaps it is victory enough to come so close to a long-standing result that was originally obtained by a delicate problem specific, argument.

Compared to Garsia's argument, the proof of (36) is straightforward. It is also reasonably robust and potentially capable of further development, even though there seems to be no room to improve the constant. In spirit the proof is close to the elegant argument of Stout (1974, pp. 145–148) for his version of the easier unweighted inequality (Proposition 2). In each instance, the heart of the matter is the application of the maximal L^2 inequality to some martingale. Here we have the benefit of ready access to the martingales (4) and (6) that were given by our general construction.

10. Observations and Connections

Our focus here is on methodology, and our primary aim has been to demonstrate the usefulness of a linear algebraic method for constructing martingales from systems of conditional expectations that are calculated for sampling without replacement. Through our examples we hope to have shown that the martingales given by our general construction have honest bite. In particular, these martingales yield reasonably direct proofs of a variety of permutation inequalities — both new ones and old ones.

Among our new inequalities, the simple Hardy-type inequality (10) seems particularly attractive. If we had to isolate a single open problem for attention, then our choice would be to determine if the constant of inequality (10) is best possible. This problem seems feasible, but one will not get any help from the easy arguments that show that the corresponding constant in Hardy's inequality is best possible.

The other new inequality that seems noteworthy is the Garsia-type inequality (8) where we introduce $V(a)$, the measure of multiplier cancelations. This inequality may be long-winded, but it isolates a common situation where one can do substantially better than the classic Garsia bound (36). The quadratic permutation inequality (Proposition 4) and the discrete bridge inequality (26) are more specialized, and they may have a hard time finding regular employment. Still, they are perfect for the right job, and they also illustrate the diversity of the martingales that come from the general construction.

We have developed several results in theory of permutation inequalities to test the effectiveness of the permutation martingales given by our construction, but permutation inequalities have a charm of their own, and one could always hope to do more. We have already mentioned the remarkable maximal inequalities of

Chobanyan (1994) and Chobanyan and Salehi (2001) that exploit combinatorial mapping arguments in addition to martingale arguments. It would be interesting to see if our new martingales could help more in that context. Finally, we did not touch on the important weak-type (or Levy-type) permutation inequalities such as those studied in Pruss (1998) and Levental (2001), but it seems reasonable to expect that the martingales (4) and (6) could also be useful in the theory of weak-type inequalities.

References

[1] Chobanyan, S., Convergence a.s. of Rearranged Random Series in Banach Space and Associated Inequalities, *Probability in Banach spaces, 9*, 3-29, Burkhauser, Boston, 1994.

[2] Chobanyan, S. and Salehi, H., Exact maximal inequalities for exchangeable systems of random variables. *Theory of Probability and its Applications*, **45**, 424-435, 2001.

[3] Garsia, A., Existence of Almost Everywhere Convergent Rearrangements for Fourier Series of L_2 functions, *Annals of Mathematics*, **79**, 623-629, 1964.

[4] Garsia, A., Combinatorial Inequalities and Convergence of Some Orthonormal Expansions, *Orthonormal Expansions and Their Continuous Analogues*, D.T. Haimo, ed., 75-98, Southern Illinois University Press, Carbondale, Illinois, 1968.

[5] Garsia, A., *Topics in almost everywhere convergence*, Markham Pub. Co., Chicago, 1970.

[6] Hardy, G.H., Note on an Inequality of Hilbert, *Math. Zeitscr.*, **6**, 314–317.

[7] Isserlis, L., On the Moment Distributions of Moments in the Case of Samples Drawn from a Limited Universe, *Proc. R. Soc. Lond. A* **132**, 586-604, 1931.

[8] Feller, W., *An Introduction to Probability Theory and Its Applications, Volume II*, John Wiley & Sons, Inc, New York, 2nd Edition, 1971.

[9] Levental, S., A Maximal Inequality for Real Numbers with Application to Exchangeable Random Variables, *Theory Probab. Appl.* **45** (3) , 525-532, 2001.

[10] Pruss, A., A maximal inequality for partial sums of finite exchangeable sequences of random variables. *Proc. Amer. Math. Soc.*, **126**, 1811-1819, 1998.

[11] Serfling, R.J., Probability Inequalities for the Sum in Sampling without Replacement, *The Annals of Statistics*, **2**, 39-48, 1974.

[12] Shiryaev, A.N., *Probability*, Springer, New York, 2nd Edition, 1995.

[13] Steele, J.M., *The Cauchy Schwarz Master Class*, Cambridge Univesity Press, Cambridge U.K., 2004.

[14] Stout, W.F., *Almost Sure Convergence*, Academic Press, Inc, New York, 1974.