THE DISTRIBUTION OF VALUES IN THE QUADRATIC ASSIGNMENT PROBLEM

Alexander Barvinok and Tamon Stephen

July 2001

Abstract. We obtain a number of results regarding the distribution of values of a quadratic function f on the set of $n \times n$ permutation matrices (identified with the symmetric group S_n) around its optimum (minimum or maximum). We estimate the fraction of permutations σ such that $f(\sigma)$ lies within a given neighborhood of the optimal value of f and relate the optimal value with the average value of f over a neighborhood of the optimal permutation. We describe a natural class of functions (which includes, for example, the objective function in the Traveling Salesman Problem) with a relative abundance of near-optimal permutations. Also, we identify a large class of functions f with the property that permutations close to the optimal permutation in the Hamming metric of S_n tend to produce near optimal values of f (such is, for example, the objective function in the symmetric Traveling Salesman Problem). We show that for general f, just the opposite behavior may take place: an average permutation in the vicinity of the optimal permutation may be much worse than an average permutation in the whole group S_n.

1. Introduction

The Quadratic Assignment Problem (QAP for short) is an optimization problem on the symmetric group S_n of $n!$ permutations of an n-element set. The QAP is one of the hardest problems of combinatorial optimization, whose special cases include the Traveling Salesman Problem (TSP) among other interesting problems.

Recently the QAP has been of interest to many people. An excellent survey is found in [Burkard *et al.* 1999]. Despite this work, it is still extremely difficult to solve QAP's of size $n = 20$ to optimality, and the solution to a QAP of size $n = 30$ is considered noteworthy, see, for example, [Anstreicher *et al.*, to appear] and [Brüngger *et al.* 1998]. Moreover, it appears that essentially no positive approximability results for the general QAP are known, although some "bad news"

1991 *Mathematics Subject Classification.* 90C27, 90C59, 68R05, 05E10.

Key words and phrases. Quadratic Assignment Problem, distribution, symmetric group, randomized algorithms, heuristics, representation theory.

This research was partially supported by NSF Grant DMS 9734138.

(non-approximability) and approximability for special classes have been established, see [Ausiello *et al.* 1999] and [Arkin *et al.* 2001].

The goal of this paper is to study the distribution of values of the objective function of the QAP. We hope that our results would allow one on one hand to understand the behavior of various heuristics, and, on the other hand, to estimate the optimum using some simple algorithms based on random or partial enumeration with guaranteed complexity bounds. In particular, we estimate how well the sample optimum from a random sample of a given size approximates the global optimum.

(1.1) The Quadratic Assignment Problem. Let Mat_n be the vector space of all real $n \times n$ matrices $A = (a_{ij})$, $1 \leq i, j \leq n$ and let S_n be the group of all permutations σ of the set $\{1, \ldots, n\}$. There is an action of S_n on the space Mat_n by simultaneous permutations of rows and columns: we let $\sigma(A) = B$, where $A = (a_{ij})$ and $B = (b_{ij})$, provided $b_{\sigma(i)\sigma(j)} = a_{ij}$ for all $i, j = 1, \ldots, n$. One can check that $(\sigma\tau)A = \sigma(\tau A)$ for any two permutations σ and τ. There is a standard scalar product on Mat_n:

$$\langle A, B \rangle = \sum_{i,j=1}^{n} a_{ij} b_{ij}, \quad \text{where} \qquad A = (a_{ij}) \quad \text{and} \quad B = (b_{ij}).$$

Let us fix two matrices $A = (a_{ij})$ and $B = (b_{ij})$ and let us consider a real-valued function $f : S_n \longrightarrow \mathbb{R}$ defined by

$$f(\sigma) = \langle B, \sigma(A) \rangle = \sum_{i,j=1}^{n} b_{\sigma(i)\sigma(j)} a_{ij} = \sum_{i,j=1}^{n} b_{ij} a_{\sigma^{-1}(i)\sigma^{-1}(j)} \tag{1.1.1}$$

The problem of finding a permutation σ where the maximum or minimum value of f is attained is known as the *Quadratic Assignment Problem.* It is one of the hardest problems of Combinatorial Optimization. From now on we assume that $n \geq 4$.

Our approach produces essentially identical results for a more general problem.

(1.2) The generalized problem. Suppose we are given a 4-dimensional array (tensor) $C = \left\{ c_{kl}^{ij} : 1 \leq i, j, k, l \leq n \right\}$ of n^4 real numbers and the function f is defined by

$$f(\sigma) = \sum_{i,j=1}^{n} c_{\sigma(i)\sigma(j)}^{ij}. \tag{1.2.1}$$

If $c_{kl}^{ij} = a_{ij} b_{kl}$ for some matrices $A = (a_{ij})$ and $B = (b_{kl})$, we get the special case (1.1.1) we started with. The convenience of working with the generalized problem is that the set of objective functions (1.2.1) is a vector space.

We introduce the standard Hamming metric on the symmetric group S_n.

(1.3) Definition. For two permutations $\tau, \sigma \in S_n$, let the distance $\mathrm{dist}(\sigma, \tau)$ be the number of indices $1 \leq i \leq n$ where σ and τ disagree:

$$\mathrm{dist}(\tau, \sigma) = |i : \sigma(i) \neq \tau(i)|.$$

One can observe that the distance is invariant under the left and right actions of S_n:

$$\mathrm{dist}(\sigma\sigma_1, \sigma\sigma_2) = \mathrm{dist}(\sigma_1, \sigma_2) = \mathrm{dist}(\sigma_1\sigma, \sigma_2\sigma)$$

for all $\sigma_1, \sigma_2, \sigma \in S_n$.

For a permutation τ and an integer $k \geq 0$, we consider the "k-th ring" around τ:

$$U(\tau, k) = \big\{\sigma \in S_n : \mathrm{dist}(\sigma, \tau) = n - k\big\}.$$

Hence for any permutation τ the group S_n splits into the disjoint union of n rings $U(\tau, k)$ for $k = 0, 1, \ldots, n-2, n$. The ring $U(\tau, k)$ consists of the permutations that agree with τ on precisely k symbols.

Let $f : S_n \longrightarrow \mathbb{R}$ be a function of type (1.1.1) or (1.2.1). Let

$$\overline{f} = \frac{1}{n!} \sum_{\sigma \in S_n} f(\sigma)$$

be the average value of f on the symmetric group and let

$$f_0 = f - \overline{f}$$

be the "shifted" function. Hence the average value of f_0 is 0. Let τ be a permutation where the maximum value of f_0 is attained, so $f_0(\tau) \geq f_0(\sigma)$ for all $\sigma \in S_n$ and $f_0(\tau) > 0$ unless $f_0 \equiv 0$ (the problem with minimum instead of maximum is completely similar). We maintain the definitions of $\overline{f}$, f_0 and τ in future sections.

We are interested in the following questions:

• Given a constant $0 < \gamma < 1$, how many permutations $\sigma \in S_n$ satisfy $f_0(\sigma) \geq \gamma f_0(\tau)$? In particular, how well the sample optimum of a set of randomly chosen permutations approximates the true optimum?

• How does the average value of f_0 over the k-th ring $U(\tau, k)$ compare with the optimal value $f_0(\tau)$? In particular, is a random permutation from the vicinity of the optimal permutation better than a random permutation from the whole group S_n?

We remark that it is easy to compute the average value $\overline{f}$ (see Lemma 6.1).

The paper is organized as follows. In Sections 2–5 we state our main results for the general QAP (Section 5) and its particular cases (Sections 2–4) and in Sections 6–11 we provide proofs. In Section 12, we formulate some open problems.

(1.4) Notation. For an integer $m \geq 0$, let

$$d_m = \sum_{j=0}^{m} (-1)^j \frac{1}{j!}.$$

For $m \geq 2$ let

$$\nu(m) = \frac{d_{m-2}}{d_m}.$$

It is convenient to agree that $\nu(0) = 0$ (and $\nu(1)$ is not defined). One can see that

$$\lim_{m \longrightarrow +\infty} \nu(m) = 1$$

and that $0 \leq \nu(m) < 1$ if m is odd and $\nu(m) > 1$ if $m \geq 2$ is even.

The following functions p and t on the symmetric group S_n play a special role in our approach. For a permutation $\sigma \in S_n$, let

$$p(\sigma) = \big| i : \sigma(i) = i \big|$$

be the number of fixed points of σ and let

$$t(\sigma) = \big| i < j : \sigma(i) = j \quad \text{and} \quad \sigma(j) = i \big|$$

be the number of 2-cycles in σ.

One can show that $p(\sigma)$, $p^2(\sigma)$ and $t(\sigma)$ are functions of type (1.2.1) for some particular tensors $\{c_{kl}^{ij}\}$, see Remark 7.6. We denote by ε the identity permutation in S_n and by $|X|$ the cardinality of a set X.

2. The Bullseye Case

Suppose that the matrix $A = (a_{ij})$ in (1.1) is symmetric and has constant row and column sums and a constant diagonal:

$$a_{ij} = a_{ji} \quad \text{for all} \quad 1 \leq i, j \leq n;$$

$$\text{for some} \quad a$$

$$\sum_{i=1}^{n} a_{ij} = a \quad \text{for all} \quad j = 1, \ldots, n \quad \text{and}$$

$$\sum_{j=1}^{n} a_{ij} = a \quad \text{for all} \quad i = 1, \ldots, n;$$

$$a_{ii} = b \quad \text{for some} \quad b \quad \text{and all} \quad i = 1, \ldots, n.$$

For example,

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 0 & \dots & 1 & 0 \end{pmatrix}, \quad a_{ij} = \begin{cases} 1 & \text{if } |i-j| = 1 \mod n \\ 0 & \text{otherwise} \end{cases}$$

satisfies these properties and the corresponding optimization problem is the *Symmetric Traveling Salesman Problem.*

Similarly, for the generalized problem (1.2) we assume that for any k and l the matrix $A = (a_{ij})$, where $a_{ij} = c_{kl}^{ij}$, is symmetric with constant row and column sums and has a constant diagonal.

It turns out that the optimum has a characteristic "bullseye" feature with respect to the averages over the rings $U(\tau, k)$ (see Definition 1.3).

(2.1) Theorem. *Let*

$$\alpha(n, k) = \frac{k^2 - 3k + \nu(n-k)}{n^2 - 3n},$$

where $k = 0, 1, \dots, n-2, n$ *and* ν *is the function of (1.4). Then we have*

$$\frac{1}{|U(\tau, k)|} \sum_{\sigma \in U(\tau, k)} f_0(\sigma) = \alpha(n, k) f_0(\tau)$$

for $k = 0, 1, \dots, n-2, n$.

(2.2) The "bullseye" distribution. We observe that as the ring $U(\tau, k)$ contracts to the optimal permutation τ (hence k increases), the *average* value of f on the ring steadily improves (as long as $k \geq 3$).

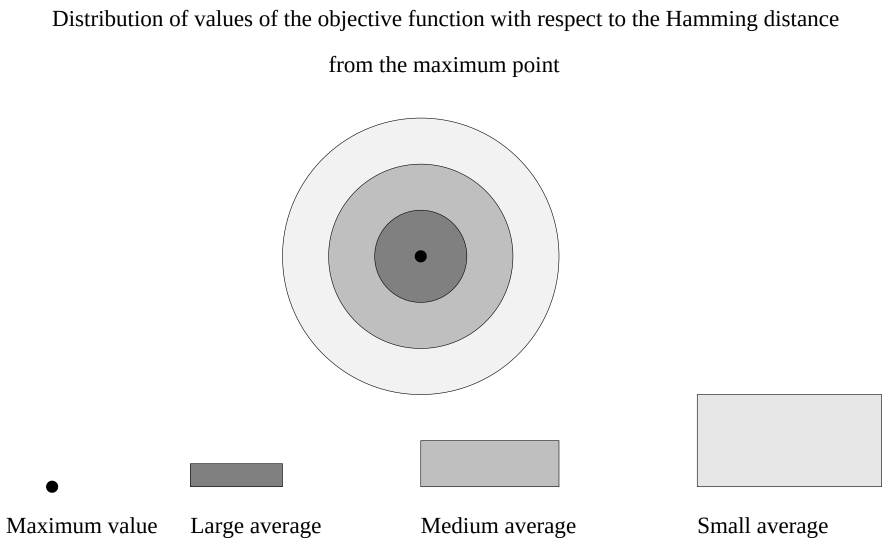

Figure 1: "Bullseye"

It is easy to construct examples where *some* values of f in a very small neighborhood of the optimum are particularly bad, but as follows from Theorem 2.1, such values are relatively rare. In our opinion, this suggests that this special case may be more amenable to various heuristics (see, for example, [Burkard *et al.* 1999]) than the general symmetric case (cf. Section 4). Incidentally, one can observe the same type of the "bullseye" behavior for the Linear Assignment Problem and some other polynomially solvable problems, such as the weighted Matching Problem.

Estimating the size of the ring $U(\tau, k)$, we get the following result.

(2.3) Theorem. *Let us choose an integer $3 \leq k \leq n-3$ and a number $0 < \gamma < 1$ and let*

$$\beta(n,k) = \frac{k^2 - 3k}{n^2 - 3n}.$$

The probability that a random permutation $\sigma \in S_n$ satisfies the inequality

$$f_0(\sigma) \geq \gamma \beta(n,k) f_0(\tau)$$

is at least

$$\frac{(1-\gamma)\beta(n,k)}{3k!}.$$

We prove our results in Section 8.

3. The Pure Case

In this section, we consider a more general case of a not necessarily symmetric matrix A in (1.1) having constant row and column sums and a constant diagonal:

$$\begin{aligned} & \text{for some} \quad a \\ \sum_{i=1}^n a_{ij} = a \quad & \text{for all} \quad j = 1, \ldots, n \quad \text{and} \\ \sum_{j=1}^n a_{ij} = a \quad & \text{for all} \quad i = 1, \ldots, n; \end{aligned}$$

$$a_{ii} = b \quad \text{for some} \quad b \quad \text{and all} \quad i = 1, \ldots, n.$$

For example, matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & 0 & 1 \\ 1 & 0 & \ldots & \ldots & \ldots & 0 \end{pmatrix}, \qquad a_{ij} = \begin{cases} 1 & \text{if } j = i+1 \mod n \\ 0 & \text{otherwise} \end{cases}$$

satisfies these properties and the corresponding optimization problem is the *Asymmetric Traveling Salesman Problem.*

Similarly, for generalized problems (1.2) we assume that for any k and l the matrix $A = (a_{ij})$, where $a_{ij} = c_{kl}^{ij}$ has constant row and column sums and has a constant diagonal.

We call this case pure, because as we remark in Sections 7 and 9, the objective function f lacks the component attributed to the Linear Assignment Problem. More generally, an arbitrary objective function f in the Quadratic Assignment Problem can be represented as a sum $f = f_1 + f_2$, where f_1 is the objective function in a Linear Assignment Problem and f_2 is the objective function in some pure case.

The behavior of averages of f_0 over the rings $U(\tau, k)$ is described by the following result.

(3.1) Theorem. *Let us define three functions of* n *and* k*:*

$$\begin{aligned}
\alpha_1(n,k) &= 1 - \nu(n-k) \\
\alpha_{2e}(n,k) &= \frac{k^2 - 3k - n - 3\nu(n-k) + \nu(n-k)n + 4}{n^2 - 4n + 4} \qquad \textit{and} \\
\alpha_{2o}(n,k) &= \frac{k^2 - 3k - n - 2\nu(n-k) + \nu(n-k)n + 3}{n^2 - 4n + 3},
\end{aligned}$$

where $k = 0, 1, \dots, n-2, n$ *and* ν *is the function of (1.4).*

If n *is even, then for some non-negative* γ_1 *and* γ_2 *such that* $\gamma_1 + \gamma_2 = 1$ *we have*

$$\frac{1}{|U(\tau,k)|} \sum_{\sigma \in U(\tau,k)} f_0(\sigma) = \Big(\gamma_1 \alpha_1(n,k) + \gamma_2 \alpha_{2e}(n,k)\Big) f_0(\tau)$$

for $k = 0, 1, \dots, n-2, n$.

If n *is odd, then for some non-negative* γ_1 *and* γ_2 *such that* $\gamma_1 + \gamma_2 = 1$ *we have*

$$\frac{1}{|U(\tau,k)|} \sum_{\sigma \in U(\tau,k)} f_0(\sigma) = \Big(\gamma_1 \alpha_1(n,k) + \gamma_2 \alpha_{2o}(n,k)\Big) f_0(\tau)$$

for $k = 0, 1, \dots, n-2, n$.

It follows from our proof (see Remark 9.3) that for any choice of $\gamma_1, \gamma_2 \geq 0$ such that $\gamma_1 + \gamma_2 = 1$ there is a function f of type (1.2.1) for which the averages of f_0 over $U(\tau, k)$ are given by the formulas of Theorem 3.1.

We observe that there are two extreme cases. If $\gamma_1 = 0$ and $\gamma_2 = 1$ then f exhibits a bullseye type distribution of Section 2.2. If $\gamma_1 = 1$ and $\gamma_2 = 0$ then f exhibits a "damped oscillator" type of distribution: the average value of f_0 over $U(\tau, k)$ changes its sign with the parity of k and approaches 0 fast as k gets smaller.

In short, if f has a damped oscillator distribution, there is no particular advantage in choosing a permutation in the vicinity of the optimal permutation τ.

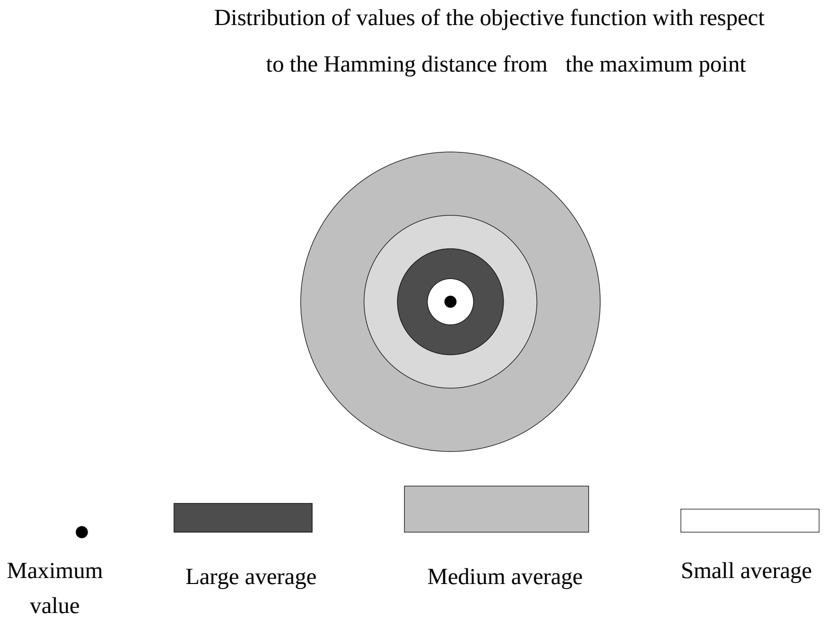

Figure 2: "Damped oscillator"

For a typical function f one can expect both γ_1 and γ_2 positive, so f would show a "weak" bullseye distribution: the average value of f_0 over $U(\tau, k)$ improves moderately as k increases, but not as dramatically as in the bullseye case of Section 2. Still, it turns out that we can find sufficiently many reasonably good permutations in the vicinity of the optimal permutation (see Remark 9.5).

(3.2) Theorem. *Let us choose an integer* $3 \leq k \leq n-3$ *and a number* $0 < \gamma < 1$ *and let*

$$\beta(n,k) = \frac{k^2 - 3k + 1}{n^2 - 3n + 1}.$$

The probability that a random permutation $\sigma \in S_n$ *satisfies the inequality*

$$f_0(\sigma) \geq \gamma \beta(n,k) f_0(\tau)$$

is at least

$$\frac{(1-\gamma)\beta(n,k)}{10k!}.$$

In particular, by choosing an appropriate k, we obtain the following corollary.

(3.3) Corollary.

(1) *Let us fix any* $\alpha > 1$. *Then there exists a* $\delta = \delta(\alpha) > 0$ *such that for all sufficiently large* $n \geq N(\alpha)$ *the probability that a random permutation* σ *in* S_n *satisfies the inequality*

$$f_0(\sigma) \geq \frac{\alpha}{n^2} f_0(\tau)$$

is at least δn^{-2}. *In particular, one can choose* $\delta = \exp\{-c\sqrt{\alpha}\ln\alpha\}$ *for some absolute constant* $c > 0$.

(2) *Let us fix any* $\epsilon > 0$. *Then there exists a* $\delta = \delta(\epsilon) < 1$ *such that for all sufficiently large* $n \geq N(\alpha)$ *the probability that a random permutation* σ *in* S_n *satisfies the inequality*

$$f_0(\sigma) \geq n^{-\epsilon} f_0(\tau)$$

is at least $\exp\{-n^{\delta}\}$. *In particular, one can choose any* $\delta > 1 - \epsilon/2$.

□

It follows from Corollary 3.3 that to get a permutation σ which satisfies (1) for any fixed α, we can use the following straightforward randomized algorithm: sample $O(n^2)$ random permutations $\sigma \in S_n$, compute the value of f and choose the best permutation. With probability tending to 1 as $n \longrightarrow +\infty$, we will hit the right permutation. The complexity of the algorithm is quadratic in n for any α, but the coefficient of n^2 grows as α grows. If we are willing to settle for an algorithm of a mildly exponential complexity of the type $\exp\{n^{\beta}\}$ for some $\beta < 1$ we can achieve a better approximation (2) by searching through the set of randomly selected $\exp\{n^{\beta}\}$ permutations. We remark that no algorithm solving the Quadratic Assignment Problem (even in the special case considered in this section) with an exponential in n complexity $\exp\{O(n)\}$ is known, although there is a dynamic programming algorithm solving the Traveling Salesman Problem in $\exp\{O(n)\}$ time.

We prove our results in Section 9.

4. The Symmetric Case

In this section, we assume that the matrix $A = (a_{ij})$ in (1.1) is symmetric, that is

$$a_{ij} = a_{ji} \quad \text{for all} \quad 1 \leq i, j \leq n.$$

Similarly, in the generalized problem (1.2) we assume that for any k and l the matrix $A = (a_{ij})$, where $a_{ij} = c_{kl}^{ij}$, is symmetric.

Overall, the distribution of values of f turns out to be much more complicated than in the special cases described in Sections 2 and 3.

(4.1) Theorem. *Let us define three functions of* n *and* k:

$$\alpha_1(n,k) = \frac{2nk - 2n - k^2 - 3k - \nu(n-k) + 6}{n^2 - 5n + 6}$$

$$\alpha_{2e}(n,k) = \frac{-nk + n + k^2 + k + \nu(n-k) - 4}{2n - 4} \quad \textit{and}$$

$$\alpha_{2o} = \frac{-n^2k + nk^2 + n^2 + nk + n\nu(n-k) - 4n - 3k + 3}{2n^2 - 7n + 3},$$

where $k = 0, 1, \dots, n-2, n$ *and* ν *is the function of (1.4).*

If n is even, then for some non-negative γ_1 and γ_2 such that $\gamma_1+\gamma_2=1$ we have

$$\frac{1}{|U(\tau,k)|}\sum_{\sigma\in U(\tau,k)} f_0(\sigma)=\Big(\gamma_1\alpha_1(n,k)+\gamma_2\alpha_{2e}(n,k)\Big)f_0(\tau)$$

for $k=0,1,\ldots,n-2,n$.

If n is odd, then for some non-negative γ_1 and γ_2 such that $\gamma_1+\gamma_2=1$ we have

$$\frac{1}{|U(\tau,k)|}\sum_{\sigma\in U(\tau,k)} f_0(\sigma)=\Big(\gamma_1\alpha_1(n,k)+\gamma_2\alpha_{2o}(n,k)\Big)f_0(\tau)$$

for $k=0,1,\ldots,n-2,n$.

It follows from our proof (see Remark 10.3) that for any choice of $\gamma_1,\gamma_2\geq 0$ such that $\gamma_1+\gamma_2=1$ there is a function f of type (1.2.1) for which that averages of f_0 over $U(\tau,k)$ are given by the formulas of Theorem 4.1. Moreover, one can choose f to be a function of type (1.1.1), but we don't prove that here (we plan to include it in [Stephen, in preparation]).

(4.2) The "spike" distribution. As in Section 3, we see that there are two extreme cases. If $\gamma_1=1$ and $\gamma_2=0$ then f has a bullseye type distribution described in Section 2.2. If $\gamma_1=0$ and $\gamma_2=1$ then f has what we call a "spike" distribution.

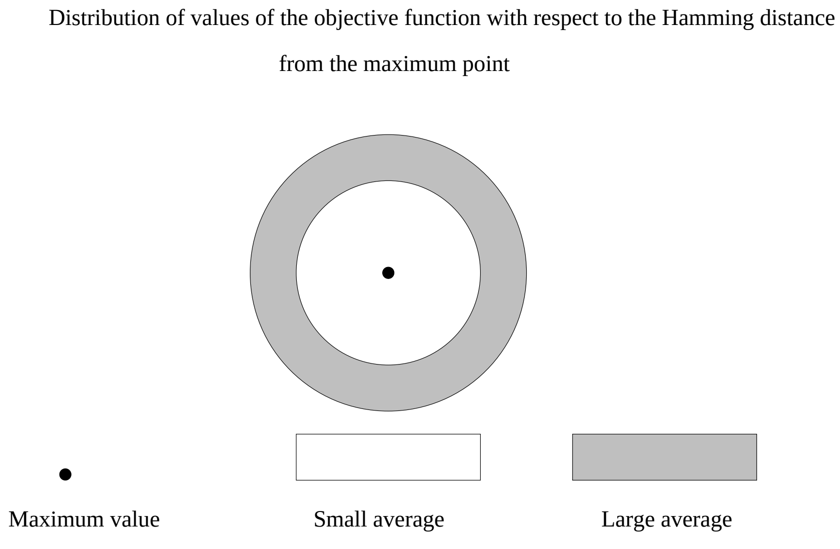

Figure 3: "Spike"

In this case, for $2\leq k\leq n-3$ the average value of f_0 over $U(\tau,k)$ is negative. Thus an average permutation $\sigma\in U(\tau,n-3)$ presents us with a worse choice than the average permutation in S_n. However, the average value of f_0 over $U(\tau,0)$ is about one half of the maximum value $f_0(\tau)$. Thus there are plenty of reasonably good permutations very far from τ and we can easily get such a permutation by random sampling.

The distribution of a typical function f around its optimal permutation in the symmetric QAP is a certain mixture of the bullseye and spike distributions. This interference of the bullseye and spike distributions (which, in some sense, are "pulling in the opposite directions") provides, in our opinion, a plausible explanation of the computational hardness of the general symmetric QAP even in comparison with other NP-hard problems such as TSP.

We obtain the following estimate for the number of near-optimal permutations.

(4.3) Theorem. *Let us choose an integer* $3 \leq k \leq n-3$ *and a number* $0 < \gamma < 1$ *and let*

$$\beta(n,k) = \frac{3k-5}{n^2 - kn + k + 2n - 5}.$$

The probability that a random permutation $\sigma \in S_n$ *satisfies the inequality*

$$f_0(\sigma) \geq \gamma \beta(n,k) f_0(\tau)$$

is at least

$$\frac{(1-\gamma)\beta(n,k)}{5k!2^k}.$$

One can notice that the obtained bound is essentially weaker than the bounds of Theorems 2.3 and 3.2. We give an example below showing that at least for the generalized problem (1.2), the bounds of Sections 2–3, and Corollary 3.3, in particular, can not hold true. The question whether the estimates can be improved for the problem (1.1) remains open.

(4.4) Scarcity of relatively good values. Let us fix any $0 < \delta < 1$ and any $0 < \epsilon < 1 - \delta$. Let us choose an integer m such that $n^{1-\epsilon} > m > n^{\delta}$. In Section 10 (cf. Remark 10.5), we show that there exists a tensor c_{kl}^{ij} such that $c_{kl}^{ij} = c_{kl}^{ji}$ for all k and l and such that for the function f defined by (1.2.1) we have

$$f(\sigma) = \frac{p^2(\sigma) - mp(\sigma) + p(\sigma) + 2t(\sigma) + m - 5}{n^2 - nm + 2n + m - 5}, \tag{4.4.1}$$

where $p(\sigma)$ is the number of fixed points in σ and $t(\sigma)$ is the number of 2-cycles in σ, see (1.4). We show that $\overline{f} = 0$ and that $f(\varepsilon) = 1$ is the maximum value of f, where ε is the identity permutation.

Then, for all sufficiently large n, the value $f(\sigma) > 2/n$ can be achieved only on permutations σ with $p(\sigma) > m$. The number of such permutations σ does not exceed $\binom{n}{m}(n-m)! = n!/m!$. That is, the probability that a random permutation σ satisfies $f(\sigma) > 2/n$ does not exceed $\exp\{-n^{\delta}\}$ for large n.

We prove the results of this section in Section 10.

5. The General Case

It appears that the difference between the general case of problems (1.1) and (1.2) and the symmetric case of Section 4 is not as substantial as the difference between the symmetric case and the special cases of Sections 2 and 3.

First, we describe the behavior of averages of f_0 over the rings $U(\tau, k)$.

(5.1) Theorem. *Let us define five functions of n and k:*

$$\alpha_1(n,k) = \frac{-k(n-k)+n-2}{n-2},$$
$$\alpha_2(n,k) = 1-\nu(n-k),$$
$$\alpha_3(n,k) = \frac{2nk-3k-2n-k^2-\nu(n-k)+6}{n^2-5n+6},$$
$$\alpha_4(n,k) = \frac{k+\nu(n-k)-2}{n-2} \quad \textit{and}$$
$$\alpha_{5o}(n,k) = \frac{-2nk+3k^2-3k+\nu(n-k)n+n-3}{n^2-2n-3},$$

where $k = 0, 1, \ldots, n-2, n$ and $\nu(n-k)$ is the function of (1.4).

If n is even, then for some non-negative $\gamma_1, \gamma_2, \gamma_3$ and γ_4 such that $\gamma_1+\gamma_2+\gamma_3+\gamma_4 = 1$ we have

$$\frac{1}{|U(\tau,k)|}\sum_{\sigma\in U(\tau,k)} f_0(\sigma)$$
$$= \Big(\gamma_1\alpha_1(n,k)+\gamma_2\alpha_2(n,k)+\gamma_3\alpha_3(n,k)+\gamma_4\alpha_4(n,k)\Big)f_0(\tau)$$

for $k = 0, 1, \ldots, n-2, n$.

If n is odd, then for some non-negative $\gamma_1, \gamma_2, \gamma_3, \gamma_4$ and γ_5 such that $\gamma_1+\gamma_2+\gamma_3+\gamma_4+\gamma_5 = 1$ we have

$$\frac{1}{|U(\tau,k)|}\sum_{\sigma\in U(\tau,k)} f_0(\sigma)$$
$$= \Big(\gamma_1\alpha_1(n,k)+\gamma_2\alpha_2(n,k)+\gamma_3\alpha_3(n,k)+\gamma_4\alpha_4(n,k)+\gamma_5\alpha_{5o}(n,k)\Big)f_0(\tau)$$

for $k = 0, 1, \ldots, n-2, n$.

It follows from our proof (see Remark 11.3) that for any choice of $\gamma_1, \ldots, \gamma_4 \geq 0$ (n even) or $\gamma_1, \ldots, \gamma_5 \geq 0$ (n odd) summing up to 1 there is a function f of type (1.2.1) for which the averages of f_0 over $U(\tau, k)$ are given by the formulas of Theorem 5.1. Moreover, one can choose f to be a function of type (1.1.1), but we don't prove that here, see [Stephen, in preparation].

As we let all but one γ equal to 0, we obtain various extreme distributions: the bullseye (when $\gamma_3 = 1$ or $\gamma_4 = 1$, cf. Section 2), damped oscillator (when $\gamma_2 = 1$, cf. Section 3) and spike (when $\gamma_1 = 1$ or $\gamma_{5o} = 1$, cf. Section 4) types. We do not get any new type of a distribution, but we can find a sharper spike than in the symmetric case.

(5.2) Example of the sharp spike. Let us consider the function

$$f = \frac{-p(n-p)+n-2}{n-2},$$

where $p(\sigma)$ is the number of fixed points of σ, see Section 1.4. In Section 11 (cf. Remark 11.3), we show that f is a function of type (1.2.1) and that $\overline{f}=0$. The maximum value of 1 is attained at the identity and at the permutations without fixed points. All other values of f are negative.

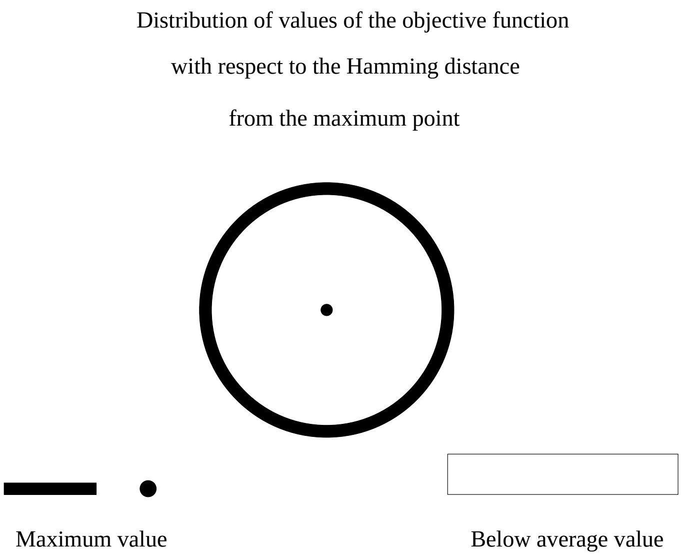

Figure 4: "Sharp spike"

One can construct a function of type (1.1.1) whose average values over the rings $U(\varepsilon, k)$, where ε is the identity permutation, coincide with those for f.

Our bound for the number of nearly optimal permutations is only slightly weaker than the bound of Theorem 4.3 in the symmetric case.

(5.3) Theorem. *Let us choose an integer* $3 \leq k \leq n-3$ *and a number* $0<\gamma<1$ *and let*

$$\beta(n,k) = \frac{k-2}{n^2-nk+k-2}.$$

The probability that a random permutation $\sigma \in S_n$ *satisfies*

$$f_0(\sigma) \geq \gamma\beta(k,n) f_0(\tau)$$

is at least

$$\frac{(1-\gamma)\beta(k,n)}{5k!}.$$

By choosing an appropriate k, we obtain the following corollary.

(5.4) Corollary.

(1) *Let us fix any $\alpha > 1$. Then there exists a $\delta = \delta(\alpha) > 0$ such that for all sufficiently large $n \geq N(\alpha)$ the probability that a random permutation σ in S_n satisfies the inequality*

$$f_0(\sigma) \geq \frac{\alpha}{n^2} f_0(\tau)$$

is at least δn^{-2}. In particular, one can choose $\delta = \exp\bigl\{-c\alpha \ln \alpha\bigr\}$ for some absolute constant $c > 0$.

(2) *Let us fix any $\epsilon > 0$. Then there exists a $\delta = \delta(\epsilon) < 1$ such that for all sufficiently large $n \geq N(\epsilon)$ the probability that a random permutation σ in S_n satisfies the inequality*

$$f_0(\sigma) \geq n^{-1-\epsilon} f_0(\tau)$$

is at least $\exp\{-n^{\delta}\}$. In particular, one can choose any $\delta > 1 - \epsilon$.

□

As in Section 2, we conclude that for any fixed $\alpha > 1$ there is a randomized algorithm of $O(n^2)$ complexity which produces a permutation σ satisfying (1). This simple randomized algorithm produces a better bound in a more general situation than some known deterministic algorithms based on semidefinite relaxations of the QAP (cf. [Ye 1999] and references therein).

If are willing to settle for an algorithm of mildly exponential complexity, we can achieve the bound of type (2), which is weaker than the corresponding bound of Corollary 3.3.

We prove our results in Section 11.

6. Preliminaries

First, we show that it is indeed easy to compute the average value $\overline{f}$ of a function f defined by (1.1.1) or (1.2.1). The result is not new (see, for example, [Graves and Whinston 1970]), we state it here for the sake of completeness.

(6.1) Lemma. *Let $f : S_n \longrightarrow \mathbb{R}$ be a function defined by (1.1.1) for some $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$. Let us define*

$$\alpha_1 = \sum_{1 \leq i \neq j \leq n} a_{ij}, \quad \alpha_2 = \sum_{i=1}^{n} a_{ii} \quad \textit{and}$$

$$\beta_1 = \sum_{1 \leq i \neq j \leq n} b_{ij}, \quad \beta_2 = \sum_{i=1}^{n} b_{ii}.$$

Then

$$\overline{f} = \frac{\alpha_1 \beta_1}{n(n-1)} + \frac{\alpha_2 \beta_2}{n}.$$

Similarly, if f is defined by (1.2.1) for some tensor $C = \left\{c_{kl}^{ij}\right\}$, $1 \leq i, j, k, l \leq n$ then

$$\overline{f} = \frac{1}{n(n-1)} \sum_{1 \leq i \neq j \leq n} \sum_{1 \leq k \neq l \leq n} c_{kl}^{ij} + \frac{1}{n} \sum_{1 \leq i, l \leq n} c_{ll}^{ii}.$$

Proof. We prove the first part only, as the proof of the the second part is completely similar. Let us choose a pair of indices $1 \leq i \neq j \leq n$. Then, as σ ranges over the symmetric group S_n, the ordered pair $\big(\sigma(i), \sigma(j)\big)$ ranges over all ordered pairs (k, l) with $1 \leq k \neq l \leq n$ and each such a pair (k, l) appears $(n-2)!$ times. Similarly, for each index $1 \leq i \leq n$, the index $\sigma(i)$ ranges over the set $\{1, \ldots, n\}$ and each $j \in \{1, \ldots, n\}$ appears $(n-1)!$ times. Therefore,

$$\begin{aligned}\overline{f} = &\frac{1}{n!} \sum_{\sigma \in S_n} \sum_{i,j=1}^{n} b_{\sigma(i)\sigma(j)} a_{ij} = \sum_{i,j=1}^{n} \Big(a_{ij} \frac{1}{n!} \sum_{\sigma \in S_n} b_{\sigma(i)\sigma(j)} \Big) \\ = &\frac{1}{n(n-1)} \sum_{i \neq j}^{n} a_{ij} \beta_1 + \frac{1}{n} \sum_{i=1}^{n} a_{ii} \beta_2 = \frac{\alpha_1 \beta_1}{n(n-1)} + \frac{\alpha_2 \beta_2}{n}\end{aligned}$$

and the proof follows. □

(6.2) Remark. Suppose that $f(\sigma) = \langle B, \sigma(A) \rangle$ for some matrices A and B and all $\sigma \in S_n$ is the objective function in the QAP (1.1) and suppose that the maximum value of f is attained at a permutation τ. Let $A_1 = \tau(A)$ and let $f_1(\sigma) = \langle B, \sigma(A_1) \rangle$. Then $f_1(\sigma) = f(\sigma\tau)$, hence the maximum value of f_1 is attained at the identity permutation ε and the distribution of values of f and f_1 is the same. We observe that if A is symmetric then A_1 is also symmetric, and if A has constant row and column sums and a constant diagonal then so does A_1 (see also Section 7). Hence, as far as the distribution of values of f is concerned, without loss of generality we may assume that the maximum of f is attained at the identity permutation ε. The same is true for functions in the generalized problem (1.2).

Next, we introduce our main tool.

(6.3) Definition. Let $f : S_n \longrightarrow \mathbb{R}$ be a function. Let us define function $g : S_n \longrightarrow \mathbb{R}$ by

$$g(\sigma) = \frac{1}{n!} \sum_{\omega \in S_n} f(\omega^{-1} \sigma \omega).$$

We call g the *central projection* of f.

It turns out that the central projection captures some important information regarding the distribution of values of a function.

(6.4) Lemma. *Let $f : S_n \longrightarrow \mathbb{R}$ be a function and let g be the central projection of f. Then*

(1) *The averages of f and g over the k-th ring $U(\varepsilon, k)$ around the identity*

permutation coincide:

$$\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)} f(\sigma)=\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)} g(\sigma);$$

(2) *The average values of f and g on the symmetric group coincide: $\overline{f}=\overline{g}$;*
(3) *Suppose that $f(\varepsilon)\geq f(\sigma)$ for all $\sigma\in S_n$. Then $g(\varepsilon)\geq g(\sigma)$ for all $\sigma\in S_n$.*

Proof. We observe that $\sigma\in U(\varepsilon,k)$ if and only if σ has exactly k fixed points. Hence for any fixed $\omega\in S_n$, the permutation $\omega^{-1}\sigma\omega$ ranges over $U(\varepsilon,k)$ as σ ranges over $U(\varepsilon,k)$. Hence

$$\begin{aligned}&\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)} g(\sigma)=\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)}\left(\frac{1}{n!}\sum_{\omega\in S_n} f(\omega^{-1}\sigma\omega)\right)\\ =&\frac{1}{n!}\sum_{\omega\in S_n}\left(\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)} f(\omega^{-1}\sigma\omega)\right)=\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma\in U(\varepsilon,k)} f(\sigma)\end{aligned}$$

and (1) is proven. Part (2) follows from (1). To prove (3), we note that $\omega^{-1}\varepsilon\omega=\varepsilon$ for all $\omega\in S_n$ and hence $g(\varepsilon)=f(\varepsilon)$. Moreover, for any $\sigma\in S_n$

$$g(\sigma)=\frac{1}{n!}\sum_{\omega\in S_n} f(\omega^{-1}\sigma\omega)\leq\frac{1}{n!}\sum_{\omega\in S_n} f(\varepsilon)=g(\varepsilon).$$

□

We will rely on a Markov type estimate, which asserts, roughly, that a function with a sufficiently large average takes sufficiently large values sufficiently often.

(6.5) Lemma. *Let X be a finite set and let $f: X\longrightarrow\mathbb{R}$ be a function. Suppose that $f(x)\leq 1$ for all $x\in X$ and that*

$$\frac{1}{|X|}\sum_{x\in X} f(x)\geq\beta\quad\text{for some}\quad\beta>0.$$

Then for any $0<\gamma<1$ we have

$$\left|\left\{x\in X: f(x)\geq\beta\gamma\right\}\right|\geq\beta(1-\gamma)|X|.$$

Proof. We have

$$\begin{aligned}\beta\leq&\frac{1}{|X|}\sum_{x\in X} f(x)=\frac{1}{|X|}\sum_{x: f(x)<\beta\gamma} f(x)+\frac{1}{|X|}\sum_{x: f(x)\geq\beta\gamma} f(x)\\ \leq&\beta\gamma+\frac{\left|\left\{x: f(x)\geq\beta\gamma\right\}\right|}{|X|}.\end{aligned}$$

Hence

$$\left|\left\{x: f(x)\geq\beta\gamma\right\}\right|\geq\beta(1-\gamma)|X|.$$

□

We need some facts about the structure of the symmetric group S_n (see, for example, [Fulton and Harris 1991]).

(6.6) The conjugacy classes of S_n

Let us fix a permutation $\rho \in S_n$. As ω ranges over the symmetric group S_n, the permutation $\omega^{-1}\rho\omega$ ranges over the conjugacy class $X(\rho)$ of ρ, that is the set of permutations that have the same cycle structure as ρ.

We will be using the following facts.

(6.6.1) Central projections and conjugacy classes. If $f : S_n \longrightarrow \mathbb{R}$ is a function and $g : S_n \longrightarrow \mathbb{R}$ its central projection, then

$$g(\rho) = \frac{1}{|X(\rho)|} \sum_{\sigma \in X(\rho)} f(\sigma).$$

If $X \subset S_n$ is a set which splits into a union of conjugacy classes $X(\rho_i) : i \in I$, and for each such a class we have

$$\frac{1}{|X(\rho_i)|} \sum_{\sigma \in X(\rho_i)} f(\sigma) \geq \alpha$$

for some number α, then

$$\frac{1}{|X|} \sum_{\sigma \in X} f(\sigma) \geq \alpha.$$

(6.6.2) Permutations with no fixed points and 2-cycles. Let us fix some positive integers $c_i : i = 1, \ldots, m$ and let a_n be the number of permutations in S_n that have no cycles of length c_i for $1 \leq i \leq m$. The exponential generating function for a_n is given by

$$\sum_{n=0}^{\infty} \frac{a_n}{n!} x^n = \frac{1}{1-x} \exp\Bigl\{ -\sum_{i=1}^{m} \frac{x^{c_i}}{c_i} \Bigr\},$$

where we agree that $a_0 = 1$, see, for example, pp. 170–173 of [Goulden and Jackson 1983]. It follows that the number of permutations $\sigma \in S_n$ without fixed points is asymptotically $e^{-1}n!$. More precisely, it is equal to $d_n n!$, see (1.4). Similarly, the number of permutations without fixed points and 2-cycles is asymptotically $e^{-3/2}n!$. We will use that the first number exceeds $n!/3$ and the second number exceeds $n!/5$ for $n \geq 3$.

We recall (1.4) that $p(\sigma)$ is the number of fixed points and $t(\sigma)$ is the number of 2-cycles of a permutation $\sigma \in S_n$. We will need a couple of technical results.

(6.7) Lemma. *We have*

$$\frac{1}{|U(\varepsilon, k)|} \sum_{\sigma \in U(\varepsilon, k)} t(\sigma) = \frac{1}{2} \nu(n-k),$$

where ν is a function of (1.4) and ε is the identity permutation.

Proof. For a pair of indices $1 \leq i < j \leq n$, let

$$t_{ij}(\sigma) = \begin{cases} 1 & \text{if } \sigma(i) = j \quad \text{and} \quad \sigma(j) = i \\ 0 & \text{otherwise.} \end{cases}$$

Then $t(\sigma) = \sum_{i<j} t_{ij}(\sigma)$. Let us compute the average of $t_{ij}(\sigma)$ over $U(\varepsilon, k)$. To choose a permutation $\sigma \in U(\varepsilon, k)$, one has to choose k fixed points in $\binom{n}{k}$ ways and then a permutation without fixed points on the remaining $n-k$ symbols in $d_{n-k}(n-k)!$ ways. Hence $|U(\varepsilon, k)| = d_{n-k} n!/k!$. To choose a permutation $\sigma \in U(\varepsilon, k)$ where (ij) is a 2-cycle, one has to choose k fixed points in $\binom{n-2}{k}$ ways and then a permutation without fixed points on $n-k-2$ symbols. Hence the total number of permutations $\sigma \in U(\varepsilon, k)$ with $t_{ij}(\sigma) = 1$ is

$$\binom{n-2}{k} d_{n-k-2}(n-k-2)! = \frac{d_{n-k-2}(n-2)!}{k!}.$$

Thus for all pairs $i < j$ we have

$$\frac{1}{|U(\varepsilon, k)|} \sum_{\sigma \in U(\varepsilon, k)} t_{ij}(\sigma) = \frac{(n-2)! d_{n-k-2}}{n! d_{n-k}} = \frac{\nu(n-k)}{n(n-1)}$$

and

$$\frac{1}{|U(\varepsilon, k)|} \sum_{\sigma \in U(\varepsilon, k)} t(\sigma) = \sum_{i<j} \frac{1}{|U(\varepsilon, k)|} \sum_{\sigma \in U(\varepsilon, k)} t_{ij}(\sigma) = \binom{n}{2} \frac{\nu(n-k)}{n(n-1)} = \frac{\nu(n-k)}{2}.$$

□

(6.8) Lemma. *For a permutation $\sigma \in S_n$, $\sigma \neq \varepsilon$, let $a_\sigma \in \mathbb{R}^2$ be the point*

$$a_\sigma = \left(p(\sigma), \quad \frac{2t(\sigma)}{n - p(\sigma)} \right).$$

Let $P = \operatorname{conv}\{a_\sigma : \sigma \neq \varepsilon\}$ be the convex hull of all such points a_σ.

If n is even, the extreme points of P are

$$(0,0), \quad (n-3, 0), \quad (n-2, 1) \quad \textit{and} \quad (0,1).$$

If n is odd, the extreme points of P are

$$(0,0), \quad (n-3,0), \quad (n-2,1), \quad \big(0, (n-3)/n\big) \quad \textit{and} \quad (1,1).$$

Proof. The set of all possible values $\big(p(\sigma), t(\sigma)\big)$, where $\sigma \neq \varepsilon$, consists of all pairs of non-negative integers (p, t) such that $p \leq n-2$, $2t \leq n$ and, additionally,

$p + 2t \le n - 3$ or $p + 2t = n$. To find the extreme points of the set of feasible points $\bigl(p, 2t/(n-p)\bigr)$, we choose a generic vector (γ_1, γ_2) and investigate for which values of p and t the maximum of

$$\gamma_1 p + \gamma_2 \frac{2t}{n-p}$$

is attained.

Clearly, we can assume that $\gamma_2 \neq 0$. If $\gamma_2 < 0$ then we should choose the smallest possible t which would be $t = 0$ unless $p = n - 2$ when we have to choose $t = 1$. Depending on the sign of γ_1, this produces the following pairs

$$(p, t) = \Bigl\{(0, 0), \quad (n-3, 0), \quad (n-2, 1)\Bigr\}.$$

If $\gamma_2 > 0$ then the largest possible value of $2t/(n-p)$ is 1. If $\gamma_1 > 0$ this produces the (already included) point

$$(p, t) = (n-2, 1\bigr).$$

If $\gamma_1 < 0$ we get

$$(p, t) = (0, n/2) \quad \text{for even } n$$

and

$$(p, t) = \Bigl\{\bigl(0, (n-3)/2\bigr), \bigl(1, (n-1)/2\bigr)\Bigr\} \quad \text{for odd} \quad n.$$

Summarizing, the extreme points of P are

$$(0, 0), \quad (n-3, 0), \quad (n-2, 1), \quad (0, 1) \qquad \text{for even } n$$

and

$$(0, 0), \quad (n-3, 0), \quad (n-2, 1), \quad \bigl(0, (n-3)/n\bigr), \quad (1, 1) \qquad \text{for odd } n$$

as claimed. □

7. Action of the Symmetric Group in the Space of Matrices

The crucial observation for our approach is that the vector space of all central projections g of functions f defined by (1.1.1) or (1.2.1) is 4-, 3-, or 2- dimensional depending on whether we consider the general case, the cases of Sections 3 and 4 or the special case of Section 2. If we require, additionally, that $\overline{f} = 0$ then the dimensions drop by 1 to 3, 2 and 1, respectively. This fact is explained by the representation theory of the symmetric group (see, for example, [Fulton and Harris 1991]). In this section, we review some facts that we need. Our notation is inspired by the generally accepted notation of the representation theory.

We describe some important invariant subspaces of the action of S_n in the space of $n \times n$ matrices Mat_n by simultaneous permutations of rows and columns. We recall that $n \ge 4$.

(7.1) Subspace L_n. Let L_n^1 be the space of constant matrices A:

$$a_{ij} = \alpha \quad \text{for some} \quad \alpha \quad \text{and all} \quad 1 \leq i, j \leq n.$$

Let L_n^2 be the subspace of scalar matrices A:

$$a_{ij} = \begin{cases} \alpha & \text{if } \ i = j \\ 0 & \text{if } \ i \neq j \end{cases} \quad \text{for some} \quad \alpha.$$

Finally, Let $L_n = L_n^1 + L_n^2$. One can observe that $\dim L_n = 2$ and that L_n is the subspace of all matrices that remain fixed under the action of S_n.

(7.2) Subspace $L_{n-1,1}$. Let $L_{n-1,1}^1$ be the subspace of matrices with identical rows and such that the sum of entries in each row is 0:

$$A = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}, \quad \text{where} \quad \alpha_1 + \ldots + \alpha_n = 0.$$

Similarly, let $L_{n-1,1}^2$ be the subspace of matrices with identical columns and such that the sum of entries in each column is 0:

$$A = \begin{pmatrix} \alpha_1 & \alpha_1 & \ldots & \alpha_1 \\ \alpha_2 & \alpha_2 & \ldots & \alpha_2 \\ \ldots & \ldots & \ldots & \ldots \\ \alpha_n & \alpha_n & \ldots & \alpha_n \end{pmatrix}, \quad \text{where} \quad \alpha_1 + \ldots + \alpha_n = 0.$$

Finally, let $L_{n-1,1}^3$ be the subspace of diagonal matrices with the zero sum of the diagonal entries:

$$A = \begin{pmatrix} \alpha_1 & 0 & \ldots & 0 & 0 \\ 0 & \alpha_2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & \alpha_n \end{pmatrix}, \quad \text{where} \quad \alpha_1 + \ldots + \alpha_n = 0.$$

Let $L_{n-1,1} = L_{n-1,1}^1 + L_{n-1,1}^2 + L_{n-1,1}^3$. One can check that the dimension of each of $L_{n-1,1}^1$, $L_{n-1,1}^2$ and $L_{n-1,1}^3$ is $n-1$ and that $\dim L_{n-1,1} = 3n-3$. Moreover, the subspaces $L_{n-1,1}^1$, $L_{n-1,1}^2$ and $L_{n-1,1}^3$ do not contain non-trivial invariant subspaces. The action of S_n in $L_{n-1,1}$, although non-trivial, is not very complicated. One can show that if $A \in L_{n-1,1} + L_n$, then the problem (1.1) of optimizing $f(\sigma)$ reduces to the Linear Assignment Problem.

(7.3) Subspace $L_{n-2,2}$. Let us define $L_{n-2,2}$ as the subspace of all *symmetric* matrices A with row and column sums equal to 0 and zero diagonal

$$\begin{aligned}
&a_{ij} = a_{ji} \quad \text{for all} \quad 1 \leq i, j \leq n; \\
&\sum_{i=1}^{n} a_{ij} = 0 \quad \text{for all} \quad j = 1, \ldots, n; \\
&\sum_{j=1}^{n} a_{ij} = 0 \quad \text{for all} \quad i = 1, \ldots, n \quad \text{and} \\
&a_{ii} = 0 \quad \text{for all} \quad i = 1, \ldots, n.
\end{aligned}$$

One can check that $L_{n-2,2}$ is an invariant subspace and that $\dim L_{n-2,2} = (n^2 - 3n)/2$. Besides, $L_{n-2,2}$ contains no non-trivial invariant subspaces.

(7.4) Subspace $L_{n-2,1,1}$. Let us define $L_{n-2,1,1}$ as the subset of all *skew symmetric* matrices A with row and column sums equal to 0:

$$\begin{aligned}
&a_{ij} = -a_{ji} \quad \text{for all} \quad 1 \leq i, j \leq n; \\
&\sum_{i=1}^{n} a_{ij} = 0 \quad \text{for all} \quad j = 1, \ldots, n \quad \text{and} \\
&\sum_{j=1}^{n} a_{ij} = 0 \quad \text{for all} \quad i = 1, \ldots, n.
\end{aligned}$$

One can check that $L_{n-2,1,1}$ is an invariant subspace and that $\dim L_{n-2,1,1} = (n^2 - 3n)/2 + 1$. Similarly, $L_{n-2,1,1}$ contains no non-trivial invariant subspaces.

One can check that $\mathrm{Mat}_n = L_n + L_{n-1,1} + L_{n-2,2} + L_{n-2,1,1}$. The importance of the subspaces (7.1)–(7.4) is explained by the fact that they are the *isotypical components* of the irreducible representations of the symmetric group in the space of matrices.

To state the main result of this section, we recall the definitions of the central projection (see Definition 6.3) and of the functions p and t (see (1.4)).

(7.5) Proposition. *For $n \times n$ matrices A and B, where $n \geq 4$, let $f : S_n \longrightarrow \mathbb{R}$ be the function defined by (1.1.1) and let $g : S_n \longrightarrow \mathbb{R}$ be the central projection of f.*

(1) *If $A \in L_n$ then g is a scalar multiple of the constant function*

$$\chi_n(\sigma) = 1 \quad \textit{for all} \quad \sigma \in S_n;$$

(2) *If $A \in L_{n-1,1}$ then g is a scalar multiple of the function*

$$\chi_{n-1,1}(\sigma) = p(\sigma) - 1 \quad \textit{for all} \quad \sigma \in S_n;$$

(3) *If $A \in L_{n-2,2}$ then g is a scalar multiple of the function*

$$\chi_{n-2,2}(\sigma) = t(\sigma) + \frac{1}{2}p^2(\sigma) - \frac{3}{2}p(\sigma) \quad \textit{for all} \quad \sigma \in S_n;$$

(4) *If $A \in L_{n-2,1,1}$ then g is a scalar multiple of the function*

$$\chi_{n-2,1,1}(\sigma) = \frac{1}{2}p^2(\sigma) - \frac{3}{2}p(\sigma) - t(\sigma) + 1 \quad \textit{for all} \quad \sigma \in S_n.$$

Proposition 7.5 follows from the representation theory of the symmetric group (see, for example, Part 1 of [Fulton and Harris 1991]). The set of all functions $f : S_n \longrightarrow \mathbb{R}$ is identified with the (real) group algebra of the symmetric group. The center of the group algebra is spanned by the characters of the irreducible representations of S_n. The basic fact that we are using here is that if f is a matrix element in an irreducible representation of the group then the central projection must be a scalar multiple of the character of that representation. The functions $\chi_n, \chi_{n-1,1}, \chi_{n-2,2}$ and $\chi_{n-1,1,1}$ are the characters of corresponding irreducible representations of S_n for $n \geq 4$ (see Lecture 4 of [Fulton and Harris 1991]). The irreducible characters are linearly independent, and, moreover orthonormal. In particular, we will use that

$$\sum_{\sigma \in S_n} \chi_{n-1,1}(\sigma) = \sum_{\sigma \in S_n} \chi_{n-2,2}(\sigma) = \sum_{\sigma \in S_n} \chi_{n-2,1,1}(\sigma) = 0,$$

hence the average value of all but the trivial character χ_n is 0.

(7.6) Remark. We note that the functions p, p^2 and t are objective functions of type (1.2.1) in some generalized QAP (1.2). Indeed, to obtain p we choose $c_{ii}^{ii} = 1$ for all $i = 1, \ldots, n$ to be the only non-zero entries of C. To obtain p^2, we choose $c_{ij}^{ij} = 1$ for $1 \leq i, j \leq n$ to be the only non-zero entries of C. To obtain t, we choose $c_{ji}^{ij} = 1$ for all $1 \leq i < j \leq n$ to be the only non-zero entries of C. Consequently, the characters χ_n, $\chi_{n-1,1}$, $\chi_{n-2,2}$ and $\chi_{n-2,1,1}$ are objective functions of type (1.2.1).

8. The Bullseye Case. Proofs

In this section, we prove Theorem 2.1 and Theorem 2.3. The proof is based on the observation that A satisfies the conditions of Section 2 if and only if $A \in L_n + L_{n-2,2}$ (see Section 7).

Proof of Theorem 2.1. Without loss of generality, we may assume that the maximum of $f_0(\sigma)$ is attained at the identity permutation ε (see Remark 6.2). Excluding the non-interesting case of $f_0 \equiv 0$, by scaling f, if necessary, we can assume that $f_0(\varepsilon) = 1$. Let g be the central projection of f_0. Then by Parts (2) and (3) of Lemma 6.4, we have $\overline{g} = 0$ and $1 = g(\varepsilon) \geq g(\sigma)$ for all $\sigma \in S_n$. Moreover, since

$A \in L_n + L_{n-2,2}$, by Parts (1) and (3) of Proposition 7.5, g must be a linear combination of the constant function χ_n and $\chi_{n-2,2}$. Since $\overline{g} = 0$, g should be proportional to $\chi_{n-2,2}$ and since $g(\varepsilon) = 1$, we have

$$g = \frac{2}{n^2 - 3n}\chi_{n-2,2} = \frac{2t + p^2 - 3p}{n^2 - 3n}.$$

Now $\sigma \in U(\varepsilon, k)$ if and only if $p(\sigma) = k$. Applying Part (1) of Lemma 6.4 and Lemma 6.7 we get

$$\frac{1}{|U(\varepsilon,k)|}\sum_{\sigma \in U(\varepsilon,k)} f_0(\sigma) = \frac{1}{|U(\varepsilon,k)|}\sum_{\sigma \in U(\varepsilon,k)} g(\sigma) = \frac{k^2 - 3k + \nu(n-k)}{n^2 - 3n}$$

and the proof follows. □

Proof of Theorem 2.3. As in the proof of Theorem 2.1, we assume that the maximum value of f_0 is equal to 1.

Let us estimate the cardinality $|U(\tau,k)| = |U(\varepsilon,k)|$. Since $\sigma \in U(\varepsilon,k)$ if and only if σ has k fixed points, to choose a $\sigma \in U(\varepsilon,k)$ one has to choose k points in $\binom{n}{k}$ ways and then choose a permutation of the remaining $n-k$ points without fixed points. Using (6.6.2), we get

$$|U(\tau,k)| \geq \binom{n}{k}(n-k)!/3 = \frac{n!}{3k!}.$$

Applying Lemma 6.5 with $\beta = \beta(n,k)$ and $X = U(\tau,k)$, from Theorem 2.1 we conclude that

$$\begin{aligned}\mathbf{P}\Big\{\sigma \in S_n : f_0(\sigma) \geq \gamma\beta(n,k)\Big\} &\geq \frac{(1-\gamma)\beta(n,k)|U(\tau,k)|}{n!} \\ &\geq \frac{(1-\gamma)\beta(n,k)}{3k!}.\end{aligned}$$

□

9. The Pure Case. Proofs

In this section, we prove Theorem 3.1 and Theorem 3.2. We observe that A satisfies the conditions of Section 3 if and only if $A \in L_n + L_{n-2,1,1} + L_{n-2,2}$ (see Section 7). As in Section 8, the L_n component contributes just a constant to f. Since the $L_{n-1,1}$ component attributed to the Linear Assignment Problem (see Section 7.2) is absent, we call this case "pure".

We choose a more convenient basis g_1 and g_2 in the vector space spanned by $\chi_{n-2,2}$ and $\chi_{n-2,1,1}$, namely:

$$g_1 = \chi_{n-2,2} + \chi_{n-2,1,1} = p^2 - 3p + 1 \quad \text{and} \quad g_2 = \chi_{n-2,1,1} - \chi_{n-2,2} = 1 - 2t.$$

(9.1) Definition. Let K_p (where p stands for "pure") be the set of all functions $g : S_n \longrightarrow \mathbb{R}$ such that $g \in \operatorname{span}\{g_1, g_2\}$, where $g_1 = p^2 - 3p + 1$ and $g_2 = 1 - 2t$ and $g(\varepsilon) \geq g(\sigma)$ for all $\sigma \in S_n$, where ε is the identity permutation. We call K_p the *central (pure) cone.*

Identifying $\operatorname{span}\{g_1, g_2\}$ with two-dimensional vector space $\mathbb{R}^2$ (plane), we see that the conditions $g(\varepsilon) \geq g(\sigma)$ define the central cone K as a convex cone in $\mathbb{R}^2$. Our goal is to find the extreme rays r_1 and r_2 of K_p, so that every function $g \in K$ can be written as a non-negative linear combination of r_1 and r_2.

(9.2) Lemma. *For $n \geq 4$ let us define the functions r_1, r_{2e} and $r_{2o} : S_n \longrightarrow \mathbb{R}$ by*

$$\begin{aligned} r_1 &= 1 - 2t, \\ r_{2e} &= \frac{p^2 - 3p - n - 6t + 2tn + 4}{n^2 - 4n + 4} \quad \textit{and} \\ r_{2o} &= \frac{p^2 - 3p - n - 4t + 2tn + 3}{n^2 - 4n + 3}. \end{aligned}$$

Then

(1) *If n is even then K_p is a 2-dimensional convex cone with the extreme rays spanned by r_1 and r_{2e};*
(2) *If n is odd then K_p is a 2-dimensional convex cone with the extreme rays spanned by r_1 and r_{2o}. Cone K_p contains r_{2e};*
(3) *If $\varepsilon \in S_n$ is the identity, then*

$$r_1(\varepsilon) = r_{2e}(\varepsilon) = r_{2o}(\varepsilon) = 1.$$

Proof. A function $g \in K_p$ can be written as a linear combination $g = \alpha_1 g_1 + \alpha_2 g_2$. Since $p(\varepsilon) = n$ and $t(\varepsilon) = 0$, we have $g(\varepsilon) = \alpha_1(n^2 - 3n + 1) + \alpha_2$. Therefore, the inequalities $g(\varepsilon) \geq g(\sigma)$ can be written as

$$\alpha_1(n^2 - 3n + 1) + \alpha_2 \geq \alpha_1\big(p^2(\sigma) - 3p(\sigma) + 1\big) + \alpha_2\big(1 - 2t(\sigma)\big),$$

which, for $\sigma \neq \varepsilon$ is equivalent to

$$\alpha_1\big(n + p(\sigma) - 3\big) + \alpha_2 \frac{2t(\sigma)}{n - p(\sigma)} \geq 0.$$

Using Lemma 6.8, we conclude that for even n, the system is equivalent to

$$\begin{aligned} &\alpha_1 \geq 0 \\ &(n-3)\alpha_1 + \alpha_2 \geq 0 \end{aligned} \tag{9.2.1}$$

and for odd n, the system is equivalent to

$$\begin{aligned} &\alpha_1 \geq 0 \\ &(n-2)\alpha_1 + \alpha_2 \geq 0. \end{aligned} \tag{9.2.2}$$

Consequently, every solution (α_1, α_2) of (9.2.1) is a non-negative linear combination of $(0,1)$ and $(1,3-n)$ and every solution of (9.2.2) is a non-negative linear combination of $(0,1)$ and $(1,2-n)$.

The functions r_1, r_{2e} and r_{2o} are obtained from $g_2, g_1+(3-n)g_2$ and $g_1+(2-n)g_2$ respectively by scaling so that the value at the identity becomes equal to 1.

Since every solution of (9.2.1) is a solution of (9.2.2), we conclude that $r_{2e} \in K_p$ for odd n as well. $\square$

(9.3) Remark. We observe that r_1 has the damped oscillator distribution corresponding to the case of $\gamma_1 = 1$ in Theorem 3.1, whereas r_{2e} and r_{2o} both have the bullseye distribution corresponding to the case of $\gamma_2 = 1$ in Theorem 3.1. If n is even, then $r_{2o} \notin K_p$. Indeed, if σ is a product of $n/2$ commuting 2-cycles, so that $p(\sigma) = 0$ and $t(\sigma) = n/2$, then $r_{2o}(\sigma) = (n^2 - 3n + 3)/(n^2 - 4n + 3) > 1 = r_{2o}(\varepsilon)$.

The central (pure) cone

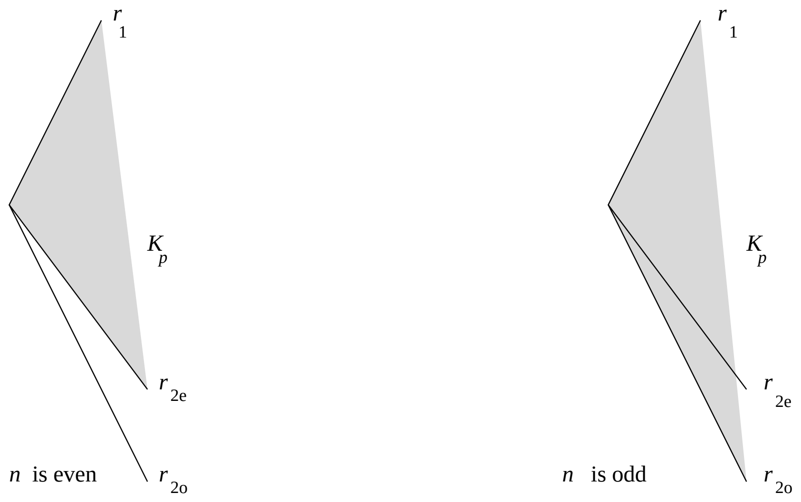

Figure 5

Proof of Theorem 3.1. We proceed as in the proof of Theorem 2.1 (Section 8) with some modifications. Without loss of generality, we assume that the maximum of $f_0(\sigma)$ is attained on the identity permutation ε and that $f_0(\varepsilon) = 1$. Let g be the central projection of f_0. Since $A \in L_n + L_{n-2,2} + L_{n-2,1,1}$, by Parts (1), (3) and (4) of Proposition 7.5, g is a linear combination of χ_n, $\chi_{n-2,2}$ and $\chi_{n-2,1,1}$. By Part (2) of Lemma 6.4, we have $\overline{g} = \overline{f_0} = 0$, so g is a linear combination of $\chi_{n-2,2}$ and $\chi_{n-2,1,1}$ alone. Moreover, by Part (3) of Lemma 6.4, we have $1 = g(\varepsilon) \geq g(\sigma)$ for all $\sigma \in S_n$. Hence g lies in the central cone K_p. Applying Lemma 9.2, we conclude that g must be a convex combination of r_1 and r_{2e} for n even and a convex combination of r_1 and r_{2o} for n odd. Applying Part (1) of Lemma 6.4, we can replace the average of f_0 over the set $U(\varepsilon, k)$ by the average of g over $U(\varepsilon, k)$. The proof now follows by Lemma 6.7 and the observation that $\sigma \in U(\varepsilon, k)$ if and only if $p(\sigma) = k$. $\square$

To prove Theorem 3.2, we need need one preliminary result.

(9.4) Lemma. *Let g be a linear combination of $g_1 = \chi_{n-2,2} + \chi_{n-2,1,1} = p^2 - 3p + 1$ and $g_2 = \chi_{n-2,1,1} - \chi_{n-2,2} = 1 - 2t$ such that $g(\varepsilon) = 1$. For a $3 \leq k \leq n-1$, let σ_k be a permutation such that $p(\sigma_k) = k$ and $t(\sigma_k) = 0$ and let θ_k be a permutation such that $p(\theta_k) = k$ and $t(\theta_k) = 1$. Then*

$$\max\{g(\sigma_k), g(\theta_k)\} \geq \frac{k^2 - 3k + 1}{n^2 - 3n + 1}.$$

Proof. Since $g(\varepsilon) = 1$, $g_1(\varepsilon) = n^2 - 3n + 1$ and $g_2(\varepsilon) = 1$, we can write

$$g = \alpha_1 \frac{p^2 - 3p + 1}{n^2 - 3n + 1} + \alpha_2(1 - 2t)$$

for some α_1 and α_2 such that $\alpha_1 + \alpha_2 = 1$. Then

$$\begin{aligned} g(\sigma_k) &= \alpha_1 \frac{k^2 - 3k + 1}{n^2 - 3n + 1} + \alpha_2 \quad \text{and} \\ g(\theta_k) &= \alpha_1 \frac{k^2 - 3k + 1}{n^2 - 3n + 1} - \alpha_2 \end{aligned}$$

We observe that $g(\sigma_k)$ and $g(\theta_k)$ are linear functions of α_1 and α_2 and that for

$$\alpha_1 = 1 \quad \text{and} \quad \alpha_2 = 0$$

we have

$$g(\sigma_k) = g(\theta_k) = \frac{k^2 - 3k + 1}{n^2 - 3n + 1} \tag{9.4.1}$$

Let

$$\lambda_1 = \frac{n^2 - 3n + 1}{2(k^2 - 3k + 1)} + \frac{1}{2} \quad \text{and} \quad \lambda_2 = \frac{n^2 - 3n + 1}{2(k^2 - 3k + 1)} - \frac{1}{2}.$$

Then $\lambda_1, \lambda_2 > 0$ and

$$\lambda_1 g(\sigma_k) + \lambda_2 g(\theta_k) = \alpha_1 + \alpha_2 = 1.$$

Comparing this with (9.4.1) we conclude that there are no values α_1 and α_2 such that $\alpha_1 + \alpha_2 = 1$ and

$$g(\sigma_k), g(\theta_k) < \frac{k^2 - 3k + 1}{n^2 - 3n + 1}.$$

□

Proof of Theorem 3.2. Without loss of generality, we may assume that the maximum value of f_0 is attained at the identity permutation ε (see Remark 6.2). Excluding an obvious case of $f_0 \equiv 0$, by scaling f, if necessary, we may assume that

$f_0(\varepsilon) = 1$. Let g be the central projection of f_0. As in the proof of Theorem 3.1, we deduce that g is a linear combination of $\chi_{n-2,2}$ and $\chi_{n-2,1,1}$ and hence a linear combination of g_1 and g_2, and that $g(\varepsilon) = 1$.

Let us choose a $3 \leq k \leq n-3$ and let X_k be the set of permutations σ such that $p(\sigma) = k$ and $t(\sigma) = 0$ and let Y_k be the set of permutations θ such that $p(\theta) = k$ and $t(\theta) = 1$. To choose a permutation $\sigma \in X_k$, one has to choose k fixed points in $\binom{n}{k}$ ways and then a permutation without fixed points or 2-cycles on the remaining $(n-k)$ points. Then, by (6.6.2)

$$|X_k| \geq \frac{1}{5}\binom{n}{k}(n-k)! = \frac{1}{5}\frac{n!}{k!}.$$

Similarly, to choose a permutation $\theta \in Y_k$, one has to choose a 2-cycle in $\binom{n}{2}$ ways, k fixed points in $\binom{n-2}{k}$ ways and a permutation without fixed points or 2-cycles on the remaining $(n-k-2)$ points. Then, by (6.6.2)

$$|Y_k| \geq \frac{1}{5}\binom{n}{2}\binom{n-2}{k}(n-k-2)! = \frac{n!}{10k!}.$$

Let us choose a permutation $\sigma \in X_k$ and a permutation $\theta \in Y_k$ and let $Z = X_k$ if $g(\sigma) \geq g(\theta)$ and $Z = Y_k$ otherwise. Then

$$|Z| \geq \frac{n!}{10k!}$$

and by Lemma 9.4,

$$g(\sigma) \geq \frac{k^2-3k+1}{n^2-3n+1} \quad \text{for all} \quad \sigma \in Z.$$

The set Z is a disjoint union of some conjugacy classes $X(\rho)$ and for each $X(\rho)$ by (6.6.1), we have

$$g(\rho) = \frac{1}{|X(\rho)|}\sum_{\sigma \in X(\rho)} f_0(\sigma) \geq \frac{k^2-3k+1}{n^2-3n+1}$$

and hence

$$\frac{1}{|Z|}\sum_{\sigma \in X(\rho)} f_0(\sigma) \geq \frac{k^2-3k+1}{n^2-3n+1}.$$

Applying Lemma 6.5 with $X = Z$ and $\beta = \beta(n,k)$, we get that

$$\mathbf{P}\Big\{\sigma \in S_n : f_0(\sigma) \geq \gamma\beta(n,k)\Big\} \geq \frac{(1-\gamma)\beta(n,k)}{10k!}.$$

$\square$

(9.5) Remark. It follows from the proof that we are able to choose the required number of "good" permutations among the permutations whose distance to the optimal permutation τ is $n-k$.

10. The Symmetric Case. Proofs

In this case, $A \in L_n + L_{n-1,1} + L_{n-2,2}$ (see Section 7). As in Sections 8 and 9, the L_n component contributes a just a constant to f. We choose a more convenient basis g_1 and g_2 in the vector space spanned by $\chi_{n-1,1}$ and $\chi_{n-2,2}$, namely

$$g_1 = \chi_{n-1,1} = p - 1 \quad \text{and} \quad g_2 = 2\chi_{n-2,2} + 3\chi_{n-1,1} = p^2 + 2t - 3,$$

where $p(\sigma)$ is the number of fixed points of σ and $t(\sigma)$ is the number of 2-cycles in σ.

(10.1) Definition. Let K_s (where s stands for "symmetric") be the set of all functions $g : S_n \longrightarrow \mathbb{R}$ such that $g \in \operatorname{span}\{g_1, g_2\}$, where $g_1 = p - 1$ and $g_2 = p^2 + 2t - 3$ and $g(\varepsilon) \geq g(\sigma)$ for all $\sigma \in S_n$, where ε is the identity permutation. We call K_s the *central (symmetric) cone.*

Identifying $\operatorname{span}\{g_1, g_2\}$ with two-dimensional vector space $\mathbb{R}^2$ (plane), we see that the conditions $g(\varepsilon) \geq g(\sigma)$ define the central cone K_s as a convex cone in $\mathbb{R}^2$. Our goal is to find the extreme rays r_1 and r_2 of K_s, so that every function $g \in K_s$ can be written as a non-negative linear combination of r_1 and r_2.

(10.2) Lemma. *For $n \geq 4$ let us define the functions r_1, r_{2e} and $r_{2o} : S_n \longrightarrow \mathbb{R}$ by*

$$\begin{aligned} r_1 &= \frac{2np - 2n - p^2 - 3p - 2t + 6}{n^2 - 5n + 6}, \\ r_{2e} &= \frac{-np + n + p^2 + p + 2t - 4}{2n - 4} \quad \textit{and} \\ r_{2o} &= \frac{-n^2p + np^2 + n^2 + np + 2nt - 4n - 3p + 3}{2n^2 - 7n + 3}. \end{aligned}$$

Then

(1) *If n is even then K_s is a 2-dimensional convex cone with the extreme rays spanned by r_1 and r_{2e};*
(2) *If n is odd then K_s is a 2-dimensional convex cone with the extreme rays spanned by r_1 and r_{2o}. Cone K_s contains r_{2e};*
(3) *If $\varepsilon \in S_n$ is the identity, then*

$$r_1(\varepsilon) = r_{2e}(\varepsilon) = r_{2o}(\varepsilon) = 1.$$

Proof. A function $g \in K_s$ can be written as a linear combination $g = \alpha_1 g_1 + \alpha_2 g_2$. Since $p(\varepsilon) = n$ and $t(\varepsilon) = 0$, we have $g(\varepsilon) = \alpha_1(n-1) + \alpha_2(n^2 - 3)$. Therefore, the inequalities $g(\varepsilon) \geq g(\sigma)$ can be written as

$$\alpha_1(n-1) + \alpha_2(n^2 - 3) \geq \alpha_1\bigl(p(\sigma) - 1\bigr) + \alpha_2\bigl(p^2(\sigma) + 2t(\sigma) - 3\bigr),$$

which, for $\sigma \neq \varepsilon$, is equivalent to

$$\alpha_1 + \alpha_2\Big(n + p(\sigma) - \frac{2t(\sigma)}{n - p(\sigma)}\Big) \geq 0. \tag{10.2.1}$$

Applying Lemma 6.8, we observe that (10.2.1) is equivalent to the system of two inequalities:

$$\alpha_1 + (2n-3)\alpha_2 \geq 0$$

and

$$\begin{aligned} &\alpha_1 + (n-1)\alpha_2 \geq 0 \quad \text{if } n \text{ is even},\\ &n\alpha_1 + (n^2 - n + 3)\alpha_2 \geq 0 \quad \text{if } n \text{ is odd}. \end{aligned}$$

Thus every pair (α_1, α_2) satisfying (10.2.1) can be written as a non-negative linear combination of $(2n-3, -1)$ and $(1-n, 1)$ when n is even and $(2n-3, -1)$ and $(-n^2+n-3, n)$ when n is odd.

The generators r_1, r_{2e} and r_{2o} are obtained from $(2n-3)g_1 - g_2$, $(1-n)g_1 + g_2$ and $(-n^2+n-3)g_1 + ng_2$ respectively by scaling so that the value at the identity becomes equal to 1.

It remains to check that $r_{2e} \in K_s$ for n odd as well. Indeed, using that $2t+p \leq n$ we have

$$\begin{aligned} (2n-4)\big(r_{2e} - 1\big) &= -n(p-1) + p(p+1) + 2t - 4 - 2n + 4\\ &= -n(p+1) + p(p+1) + 2t\\ &\leq (p+1)(-n+p) + n - p = p(-n+p) \leq 0. \end{aligned}$$

□

(10.3) Remark. We observe that r_1 has the bullseye distribution corresponding to the case of $\gamma_1 = 1$ in Theorem 4.1, whereas r_{2e} and r_{2o} both have the spike distribution corresponding to the case of $\gamma_2 = 1$ in Theorem 4.1. One can observe that if n is even then $r_{2o} \notin K_s$. Indeed, if σ is a product of $n/2$ commuting 2-cycles then $r_{2o}(\sigma) = (2n^2 - 4n + 3)/(2n^2 - 7n + 3) > 1$. The picture of K_s is very similar to that of K_p, see Section 9, Figure 5.

Proof of Theorem 4.1. The proof is completely similar to the proof of Theorem 3.1, see Section 9. We use Lemma 10.2 instead of Lemma 9.2. □

To prove Theorem 4.3, we need one preliminary result.

(10.4) Lemma. *Let g be a linear combination of $g_1 = \chi_{n-1,1} = p - 1$ and $g_2 = 2\chi_{n-2,2} + 3\chi_{n-1,1} = p^2 + 2t - 3$ such that $g(\varepsilon) = 1$. For a $3 \leq k \leq n-3$, let σ_k be a permutation such that $p(\sigma_k) = k$ and $t(\sigma_k) = 0$ and let η_k be a permutation such that $p(\eta_k) = 0$ and $t(\eta_k) = k$. Then*

$$\max\big\{g(\sigma_k), g(\eta_k)\big\} \geq \frac{3k-5}{n^2 - kn + k + 2n - 5}.$$

Proof. We can write
$$g = \alpha_1 \frac{p-1}{n-1} + \alpha_2 \frac{p^2+2t-3}{n^2-3}$$
for some α_1 and α_2 such that $\alpha_1 + \alpha_2 = 1$. Then
$$g(\sigma_k) = \alpha_1 \frac{k-1}{n-1} + \alpha_2 \frac{k^2-3}{n^2-3} \quad \text{and}$$
$$g(\eta_k) = -\alpha_1 \frac{1}{n-1} + \alpha_2 \frac{2k-3}{n^2-3}.$$
We observe $g(\sigma_k)$ and $g(\eta_k)$ are linear functions of α_1 and α_2 and that for
$$\alpha_1 = \frac{kn-k-2n+2}{-n^2+kn-k-2n+5} \quad \text{and} \quad \alpha_2 = \frac{3-n^2}{-n^2+kn-k-2n+5}$$
we have

(10.4.1) $$g(\sigma_k) = g(\eta_k) = \frac{3k-5}{n^2-kn+k+2n-5} \quad \text{and} \quad \alpha_1+\alpha_2 = 1.$$

Let
$$\lambda_1 = \frac{n^2+2kn-3n-2k}{k(3k-5)} \quad \text{and} \quad \lambda_2 = \frac{kn^2-k^2n-n^2+k^2+3n-3k}{k(3k-5)}.$$
Then $\lambda_1, \lambda_2 > 0$ and
$$\lambda_1 g(\sigma_k) + \lambda_2 g(\eta_k) = \alpha_1 + \alpha_2 = 1.$$
Comparing this with (10.4.1), we conclude that there are no values α_1, α_2 such that $\alpha_1 + \alpha_2 = 1$ and
$$g(\sigma_k), g(\eta_k) < \frac{3k-5}{n^2-kn+k+2n-5}.$$
□

Proof of Theorem 4.3. The proof is similar to that of Theorem 3.2 (Section 9). Without loss of generality, we assume that the maximum value of f_0 is 1 and is attained at the identity permutation ε. Let g be the central projection of f_0. We deduce that g is a linear combination of $\chi_{n-1,1}$ and $\chi_{n-2,2}$ and hence a linear combination of g_1 and g_2, and that $g(\varepsilon) = 1$. Let us choose a $3 \leq k \leq n-3$ and let X_k be the set of permutations σ such that $p(\sigma) = k$ and $t(\sigma) = 0$ and let Y_k be the set of permutations η such that $p(\eta) = 0$ and $t(\eta) = k$. As in the proof of Theorem 3.2, we have
$$|X_k| \geq \frac{1}{5}\frac{n!}{k!}.$$

To choose a permutation $\eta \in Y_k$, one has to choose k 2-cycles in $\dfrac{n!}{(n-2k)!k!2^k}$ ways and a permutation without fixed points or 2-cycles on the remaining $n-2k$ points. Hence, by (6.6.2)

$$|Y_k| \geq \frac{1}{5}\frac{n!}{k!2^k}.$$

Let us choose a permutation $\sigma \in X_k$ and a permutation $\eta \in Y_k$ and let $Z = X_k$ if $g(\sigma) \geq g(\eta)$ and let $Z = Y_k$ otherwise. Then

$$|Z| \geq \frac{n!}{5k!2^k}$$

and by Lemma 10.4,

$$g(\sigma) \geq \frac{3k-5}{n^2 - kn + k + 2n - 5} \quad \text{for all} \quad \sigma \in Z.$$

The proof now proceed as in the proof of Theorem 3.2, Section 9. □

(10.5) Remark. Scarcity of relatively good values. Let us consider the function f of Example 4.4. We observe that

$$f = \alpha_1 r_1 + \alpha_2 r_{2e},$$

where r_1 and r_{2e} are the functions of Lemma 10.2, for

$$\alpha_1 = \frac{n^2 - nm - 2n + 3m - 3}{n^2 - nm + 2n + m - 5} \quad \text{and} \quad \alpha_2 = \frac{4n - 2m - 2}{n^2 - nm + 2n + m - 5}.$$

Thus f is a convex combination of r_1 and r_{2e}, hence $1 = f(\varepsilon) \geq f(\sigma)$ for all $\sigma \in S_n$ and $\overline{f} = 0$. Remark 7.6 implies that f is a generalized function (1.2.1) of the required type.

11. The General Case. Proofs

Let us choose a convenient basis in $\operatorname{span}\{\chi_{n-1,1}, \chi_{n-2,2}, \chi_{n-2,1,1}\}$:

$$g_1 = \chi_{n-1,1} = p - 1, \quad g_2 = \chi_{n-2,2} + \chi_{n-2,1,1} + 3\chi_{n-1,1} = p^2 - 2 \quad \text{and}$$

$$g_3 = \chi_{n-2,1,1} - \chi_{n-2,2} = 1 - 2t.$$

(11.1) Definition. Let K be the set of all functions $g \in \operatorname{span}\{g_1, g_2, g_3\}$ such that $g(\varepsilon) \geq g(\sigma)$ for all $\sigma \in S_n$. We call K the *central cone*.

Identifying $\operatorname{span}\{g_1, g_2, g_3\}$ with a 3-dimensional vector space $\mathbb{R}^3$, we see that conditions $g(\varepsilon) \geq g(\sigma)$ define the central cone K as a convex polyhedral cone in $\mathbb{R}^3$. The condition $g(\varepsilon) = 1$ defines a plane H in $\mathbb{R}^3$ and the intersection $B = H \cap K$ is a *base* of K, that is, a polygon such that every $g \in K$ can be uniquely represented in the form $g = \lambda h$ for some $h \in B$.

Our goal is to determine the structure of K. This is somewhat more complicated than in the 2-dimensional situations of Sections 9-10.

(11.2) Lemma. *Let us define functions*

$$\begin{aligned}
r_1 &= \frac{-np+n+p^2-2}{n-2},\\
r_2 &= 1-2t,\\
r_3 &= \frac{2np-3p-2n-p^2-2t+6}{n^2-5n+6},\\
r_4 &= \frac{p+2t-2}{n-2} \qquad\qquad \textit{and}\\
r_{5o} &= \frac{-2np+3p^2-3p+2tn+n-3}{n^2-2n-3}.
\end{aligned}$$

Then

(1) *If $\varepsilon \in S_n$ is the identity, then*

$$r_1(\varepsilon) = r_2(\varepsilon) = r_3(\varepsilon) = r_4(\varepsilon) = r_{5o}(\varepsilon) = 1;$$

(2) *If n is even then r_1, r_2, r_3 and r_4 are the vertices (in consecutive order) of the planar quadrilateral $B = \operatorname{conv}\{r_1, r_2, r_3, r_4\}$ that is a base of the central cone K;*

(3) *If n is odd then r_1, r_2, r_3, r_4 and r_{5o} are the vertices (in consecutive order) of the planar pentagon $B = \operatorname{conv}\{r_1, r_2, r_3, r_4, r_{5o}\}$ that is a base of the central cone K.*

Proof. A function $g \in \operatorname{span}\{g_1, g_2, g_3\}$ can be written as a linear combination $g = \alpha_1 g_1 + \alpha_2 g_2 + \alpha_3 g_3$. Then $g(\varepsilon) = \alpha_1(n-1) + \alpha_2(n^2-2) - \alpha_3$ and the conditions $g(\varepsilon) \geq g(\sigma)$ are written as

$$\alpha_1(n-1) + \alpha_2(n^2-2) + \alpha_3 \geq \alpha_1\big(p(\sigma)-1\big) + \alpha_2\big(p^2(\sigma)-2\big) + \alpha_3\big(1-2t(\sigma)\big),$$

which for $\sigma \neq \varepsilon$ are equivalent to

$$\alpha_1 + \alpha_2\big(n+p(\sigma)\big) + \alpha_3\frac{2t(\sigma)}{n-p(\sigma)} \geq 0.$$

Applying Lemma 6.8, we see that for even n, the system is equivalent to

$$\begin{aligned}
&\alpha_1 + n\alpha_2 \geq 0\\
&\alpha_1 + (2n-3)\alpha_2 \geq 0\\
&\alpha_1 + (2n-2)\alpha_2 + \alpha_3 \geq 0\\
&\alpha_1 + n\alpha_2 + \alpha_3 \geq 0
\end{aligned} \tag{11.2.1}$$

whereas for odd n, the system is equivalent to

$$\begin{aligned}
&\alpha_1 + n\alpha_2 \geq 0\\
&\alpha_1 + (2n-3)\alpha_2 \geq 0\\
&\alpha_1 + (2n-2)\alpha_2 + \alpha_3 \geq 0\\
&\alpha_1 + (n+1)\alpha_2 + \alpha_3 \geq 0\\
&n\alpha_1 + n^2\alpha_2 + (n-3)\alpha_3 \geq 0.
\end{aligned} \tag{11.2.2}$$

The set of all feasible 3-tuples $(\alpha_1, \alpha_2, \alpha_3)$ is a polyhedral cone, which, for even n, has at most 4 extreme rays and for odd n has at most 5 extreme rays. We call an inequality of (11.2.1)–(11.2.2) *active* on a particular tuple if it holds with equality.

It is readily verified that for even n the following tuples span the extreme rays of the set of solutions to (11.2.1):

$$\begin{array}{rl}
\big(-n,\quad 1,\quad 0\big) & \text{4th and 1st inequalities are active}\\
\big(0,\quad 0,\quad 1\big) & \text{1st and 2nd inequalities are active}\\
\big(2n-3,\quad -1,\quad 1\big) & \text{2nd and 3d inequalities are active}\\
\big(1,\quad 0,\quad -1\big) & \text{3d and 4th inequalities are active}
\end{array}$$

and that for odd n the following tuples span the extreme rays of the set of solutions to (11.2.1):

$$\begin{array}{rl}
\big(-n,\quad 1,\quad 0\big) & \text{5th and 1st inequalities are active}\\
\big(0,\quad 0,\quad 1\big) & \text{1st and 2nd inequalities are active}\\
\big(2n-3,\quad -1,\quad 1\big) & \text{2nd and 3d inequalities are active}\\
\big(1,\quad 0,\quad -1\big) & \text{3d and 4th inequalities are active}\\
\big(-2n-3,\quad 3,\quad -n\big) & \text{4th and 5th inequalities are active}
\end{array}$$

We obtain r_1, r_2, r_3, r_4 and r_{5o} by scaling the corresponding linear combinations $\alpha_1 g_1 + \alpha_2 g_2 + \alpha_3 g_3$ so that the value at the identity is equal to 1 and hence r_1, r_2, r_3, r_4 and r_{5o} lie on the same plane in $\operatorname{span}\{g_1, g_2, g_3\}$. □

(11.3) Remark. We observe that r_1 and r_{5o} have spike distributions (in particular, r_1 has the "sharp spike" distribution of Section 5.2) corresponding to the cases of $\gamma_1 = 1$ and $\gamma_5 = 1$ respectively in Theorem 5.1, that r_2 has the damped oscillator distribution corresponding to the case of $\gamma_2 = 1$ and that r_3 and r_4 have bullseye distributions corresponding to the cases of $\gamma_3 = 1$ and $\gamma_4 = 1$ respectively. If n is even then $r_{5o} \notin K$, for if σ is a product of $n/2$ commuting 2-cycles, so that $p(\sigma) = 0$

and $t(\sigma)=n/2$, then $r_{5o}(\sigma)=(n^2+n-3)/(n^2-2n-3)>1=r_{5o}(\varepsilon)$.

The base of the central cone

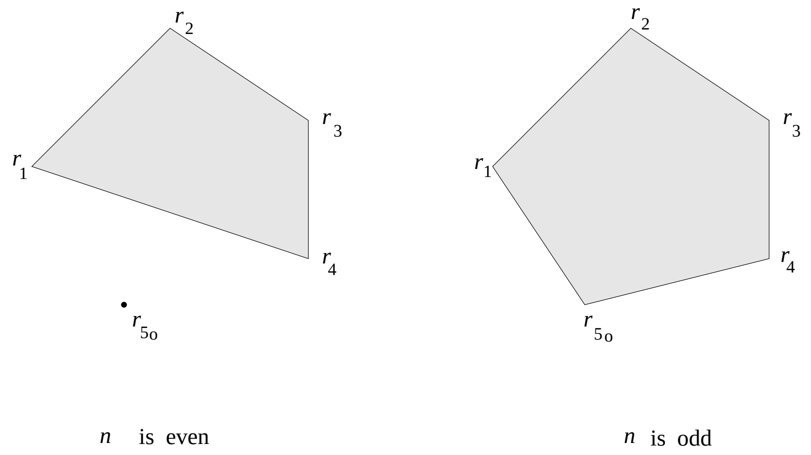

Figure 6

We observe that function r_3 coincides with function r_1 of Lemma 10.2 (the symmetric QAP) and that function r_2 coincides with function r_1 of Lemma 9.3 (the pure QAP).

Proof of Theorem 5.1. The proof is completely similar to the proof of Theorem 3.1, Section 9. We use Lemma 11.2 instead of Lemma 9.2. □

To prove Theorem 5.2, we need one preliminary result.

(11.4) Lemma. *Let g be a linear combination of $g_1=p-1$, $g_2=p^2-2$ and $g_3=1-2t$ such that such that $g(\varepsilon)=1$. For a $2\leq k\leq n-2$, let σ_k be a permutation such that $p(\sigma_k)=k$ and $t(\sigma_k)=0$, let η be a permutation such that $p(\eta)=0$ and $t(\eta)=1$ and let θ be permutation such that $p(\theta)=t(\theta)=0$. Then*

$$\max\{g(\sigma_k),g(\eta),g(\theta)\}\geq\frac{k-2}{n^2-kn+k-2}.$$

Proof. We can write

$$g=\alpha_1\frac{p-1}{n-1}+\alpha_2\frac{p^2-2}{n^2-2}+\alpha_3(1-2t)$$

for some α_1,α_2 and α_3 such that $\alpha_1+\alpha_2+\alpha_3=1$. Then

$$\begin{aligned}
g(\sigma_k)&=\alpha_1\frac{k-1}{n-1}+\alpha_2\frac{k^2-2}{n^2-2}+\alpha_3\\
g(\eta)&=-\frac{\alpha_1}{n-1}-\alpha_2\frac{2}{n^2-2}-\alpha_3\\
g(\theta)&=-\frac{\alpha_1}{n-1}-\alpha_2\frac{2}{n^2-2}+\alpha_3.
\end{aligned}$$

We observe that $g(\sigma_k)$, $g(\eta)$ and $g(\theta)$ are linear functions of α_1, α_2 and α_3 and that for

$$\alpha_1 = \frac{k(1-n)}{n^2 - nk + k - 2}, \quad \alpha_2 = \frac{n^2 - 2}{n^2 - nk + k - 2} \quad \text{and} \quad \alpha_3 = 0$$

we have

$$g(\sigma_k) = g(\eta) = g(\theta) = \frac{k-2}{n^2 - nk + k - 2} \quad \text{and} \quad \alpha_1 + \alpha_2 + \alpha_3 = 1. \tag{11.4.1}$$

Let

$$\lambda_1 = \frac{n^2 - 2n}{k^2 - 2k}, \quad \lambda_2 = \frac{n^2 - nk}{2k - 4} \quad \text{and} \quad \lambda_3 = \frac{n^2 - kn - 2n + 2k}{2k}.$$

Then $\lambda_1, \lambda_2, \lambda_3 > 0$ and

$$\lambda_1 g(\sigma_k) + \lambda_2 g(\eta) + \lambda_3 g(\theta) = \alpha_1 + \alpha_2 + \alpha_3 = 1.$$

Comparing this with (11.4.1), we conclude that there are no values α_1, α_2 and α_3 such that $\alpha_1 + \alpha_2 + \alpha_3 = 1$ and

$$g(\sigma_k), g(\eta), g(\theta) < \frac{k-2}{n^2 - nk + k - 2}.$$

□

Proof of Theorem 5.2. The proof follows those of Theorem 3.2 (Section 9) and Theorem 4.2 (Section 10) with some modifications. Let X_k be the set of all permutations σ such that $p(\sigma) = k$ and $t(\sigma) = 0$. As in the proof of Theorem 3.2, we have

$$|X_k| \geq \frac{1}{5} \frac{n!}{k!}.$$

Let Y be the set of all permutations σ such that $p(\sigma) = 0$ and $t(\sigma) = 1$. To choose a permutation $\sigma \in Y$, one has to choose a 2-cycle in $\binom{n}{2}$ ways and then an arbitrary permutation of the remaining $(n-2)$ symbols without fixed points and 2-cycles. Using (6.6.2), we estimate

$$|Y| \geq \frac{1}{5} \frac{n!}{2(n-2)!}(n-2)! = \frac{1}{10} n!.$$

Let us choose a permutation $\sigma_k \in X_k$, a permutation $\eta \in Y$ and a permutation $\theta \in X_0$. Let us choose Z to be one of X_k, X_0 and Y, depending where the maximum value of $g(\sigma_k)$, $g(\eta)$ or $g(\theta)$ is attained. Hence

$$|Z| \geq \frac{n!}{5k!}$$

and by Lemma 11.4,

$$g(\sigma) \geq \frac{k-2}{n^2 - kn + k - 2} \quad \text{for all} \quad \sigma \in Z.$$

We proceed now as in the proof of Theorem 3.2, Section 9. □

12. Concluding Remarks and Open Questions

The estimates of Theorems 2.3, 3.2, 4.3 and 5.3 for the number of near-optimal permutations can be used to bound the optimal value by a sample optimum in branch-and-bound algorithms. Those estimates are (nearly) best possible for the generalized problem (1.2). However, it is not clear whether they can be improved in the case of standard QAP (1.1) or how to improve them in interesting special cases. In particular, we ask the following

• Question: Let $f: S_n \longrightarrow \mathbb{R}$ be the objective function in the Traveling Salesman Problem (cf. Sections 2 and 3), let $\overline{f}$ be the average value of f and let $f_0 = f - \overline{f}$. Let τ be an optimal permutation, so that $f_0(\tau) \geq f_0(\sigma)$ for all $\sigma \in S_n$. Is it true that for any fixed $\gamma > 0$ there is a number $\delta = \delta(\gamma) > 0$ such that the probability that a random permutation $\sigma \in S_n$ satisfies the inequality $f_0(\sigma) \geq \dfrac{\gamma}{n} f_0(\tau)$ is at least $n^{-\delta}$ for all sufficiently large n?

Another question is whether approximations can be obtained *deterministically*. The random sampling algorithm in the "bullseye" case of Section 2 can be relatively easily derandomized. Whether the same is true for algorithms of Sections 3–5 is not clear at the moment.

It appears that the problems without linear part (the "bullseye" and "pure" cases of Sections 2 and 3) are easier than the general QAP. A possible way to use this is to split the problem into the pure quadratic and linear parts and try to "estimate away" the linear part.

Our methods can be applied to study the distribution of values in the Assignment Problems of higher order and their special cases, such as the Weighted Hypergraph Matching Problem.

References

K. Anstreicher, N. Brixius, J.-P. Goux and J. Linderoth, Solving large quadratic assignment problems on computational grids, *Math. Programming B*, to appear.

E. Arkin, R. Hassin and M. Sviridenko, Approximating the maximum quadratic assignment problem, *Inform. Process. Lett.*, **77** (2001), no. 1, 13–16.

G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi, *Complexity and Approximation. Combinatorial optimization problems and their approximability properties*, Springer-Verlag, Berlin, 1999.

A. Brüngger, A. Marzetta, J. Clausen and M. Perregaard, Solving large scale quadratic assignment problems in parallel with the search library ZRAM, *Journal of Parallel and Distributed Computing*, **50**, pp. 157-66, 1998.

R. Burkard, E. Çela, P. Pardalos and L. Pitsoulis, The quadratic assignment problem, in: *Handbook of Combinatorial Optimization* (D.-Z. Du and P.M. Pardalos, eds.), Kluwer Academic Publishers, pp. 75-149, 1999.

W. Fulton and J. Harris, *Representation Theory*, Springer-Verlag, New York, 1991.

I.P. Goulden and D.M. Jackson, Combinatorial Enumeration, *Wiley-Interscience Series in Discrete Mathematics*, John Wiley & Sons, Inc., New York, 1983.

G.W. Graves and A.B. Whinston, An algorithm for the quadratic assignment problem, *Management Science*, **17**(1970), np. 7, 453–471.

T. Stephen, The distribution of values in combinatorial optimization problems, *Ph.D. Dissertation*, Univ. of Michigan, in preparation.

Y. Ye, Approximating global quadratic optimization with convex quadratic constraints, *Journal of Global Optimization*, **15**(1999), no. 1, 1–17.

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109
E-mail address: barvinok@umich.edu, tamon@umich.edu