Chapter 6

Fourier Sampling Coset States of Nonabelian Groups

In this chapter we analyze the HSP algorithm over nonabelian groups. The motivation is a reduction of graph isomorphism, a long standing open problem, to the HSP over the symmetric group S_n. In particular, we will analyze Fourier sampling coset states when the group is nonabelian. We show that Fourier sampling a polynomial number of coset states is enough to reconstruct hidden normal subgroups. This is just a generalization of the abelian case, since all subgroups of abelian groups are normal. There are two issues that we cannot address in general as we did in the abelian case. One is whether or not the quantum Fourier transform can be computed efficiently, and the other is whether or not we can efficiently compute the intersections of kernels of the group homomorphisms. These two questions must be answered on a group by group basis. We are basically analyzing what the Fourier transform can accomplish as a change of basis. We also show that Fourier

sampling a polynomial number of coset states is not enough to distinguish trivial from nontrivial subgroups of the symmetric group, a problem motivated by the graph isomorphism problem.

The algorithm for finding normal subgroups is as follows. Recall that M_i is a measurement of register i.

Algorithm 6.1

1. *Repeat the following* $k = poly(\log(|G|))$ *times:*

a. *Create a random coset state* $|cH\rangle$*:*

$$|0\rangle \xrightarrow{F_G} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle \xrightarrow{U_f} \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, f(g)\rangle \xrightarrow{M_2} \frac{1}{\sqrt{|H|}} \sum_{h \in H} |ch\rangle$$

b. *Fourier sample* $|cH\rangle$ *and get a group homomorphism* ρ*. Call this distribution* $\mathcal{D}$*.*

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |ch\rangle \xrightarrow{F_G} \frac{1}{\sqrt{|H^\perp|}} \sum_{\rho \in H^\perp} \sum_{i,j=1}^{d_\rho} |\rho, i, j\rangle \xrightarrow{M_1} |\rho\rangle$$

2. *Classically compute the intersection:* $H = \bigcap_{i=1}^{k} \ker \rho_i$*.*

To define the Fourier transform over nonabelian groups we need some representation theory of groups. The rest of the chapter is organized as follows. In Section 6.1 some background on the representations of groups is given. In Section 6.2 a general formula is given for the distribution over representations when Fourier sampling coset states. In Section 6.3 it is shown that Fourier sampling coset states is enough to find normal hidden subgroups. In Section 6.4 it is shown that Fourier sampling a polynomial number of coset states is not enough to distinguish trivial from nontrivial subgroups in the symmetric group.

6.1 Representation Theory Background

The main tool used by polynomial time quantum algorithms is the Fourier transform. To define the Fourier transform (over a group) we require the basic elements of representation theory, defined below. See [34, 47] for more details.

Representation. A representation ρ of a group G is a homomorphism $\rho : G \to GL(V)$, where V is vector space over $\mathbb{C}$. Fixing a basis for V, each $\rho(g)$ may be realized as a $d \times d$ matrix over $\mathbb{C}$, where d is the dimension of V. As ρ is a homomorphism, for any $g, h \in G$, $\rho(gh) = \rho(g)\rho(h)$. The *dimension* d_ρ of the representation ρ is d, the dimension of V.

A representation provides a means for investigating a group by homomorphically mapping it into a family of matrices. With this realization, the group operation is matrix multiplication and tools from linear algebra can be applied to study the group. We shall be concerned with complex-valued functions on a group G; the representations of the group are relevant to this study, as they give rise to the Fourier transform for such functions.

Irreducibility. We say that a subspace W is an *invariant* subspace of a representation ρ if $\rho(g)W \subseteq W$ for all $g \in G$. We assume, without loss of generality, that every $\rho(g)$ is unitary, and, in particular, diagonalizable. Hence there are many subspaces fixed by an individual matrix $\rho(g)$. In order for W to be an invariant subspace for ρ, it must be simultaneously fixed under all $\rho(g)$.

The zero subspace and the subspace V are always invariant. If no nonzero proper

subspaces are invariant, the representation is said to be *irreducible*.

Decomposition. When a representation *does* have a nonzero proper invariant subspace $V_1 \subset V$, it is always possible to find a complementary subspace V_2 (so that $V = V_1 \oplus V_2$) which is also invariant. Since $\rho(g)$ fixes V_1, we may let $\rho_1(g)$ be the linear map on V_1 given by $\rho(g)$. It is not hard to see that $\rho_1 : G \to \mathrm{GL}(V_1)$ is in fact a representation. Similarly, define $\rho_2(g)$ to be $\rho(g)$ restricted to V_2. Since $V = V_1 \oplus V_2$, the linear map $\rho(g)$ is completely determined by $\rho_1(g)$ and $\rho_2(g)$, and in this case we write $\rho = \rho_1 \oplus \rho_2$. In this case there is a basis for V so that each matrix $\rho(g)$ is block diagonal with two blocks.

Complete Reducibility. Repeating the process described above, any representation ρ may be written as $\rho = \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$, where each ρ_i is irreducible. In particular, there is a basis in which every matrix $\rho(g)$ is block diagonal, the ith block corresponding to the ith representation in the decomposition.

Characters. The *character* $\chi_\rho : G \to \mathbb{C}$ of a representation ρ is defined by $\chi_\rho(g) = \mathrm{Tr}(\rho(g))$, the trace of the map. It is independent of the basis, and, as it turns out, completely determines the representation ρ.

Orthogonality of Characters. For two functions f_1 and f_2 on a group G, there is a natural inner product: $\langle f_1, f_2 \rangle_G$ given by $\frac{1}{|G|} \sum_g f_1(g) f_2(g)^*$. The useful fact is the following: given the character χ_ρ of any representation ρ and the character χ_i of any irreducible representation ρ_i, the inner product $\langle \chi_\rho, \chi_i \rangle$ is precisely the number of times the representation ρ_i appears in the decomposition of ρ. Since each ρ is unitary,

the inner product of two characters simplifies slightly:

$$\langle \chi_\rho, \chi_i \rangle_G = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \chi_i(g^{-1}).$$

Restriction. A representation ρ of a group G is also automatically a representation of any subgroup H. We refer to this *restricted* representation on H as $\mathrm{Res}_H \rho$. Note that even representations which are irreducible over G may be reducible when restricted to H.

Up to isomorphism, a finite group has a finite number of irreducible representations. For a group G, we let $\hat{G}$ denote this finite collection of irreducible representations. As mentioned above, any representation may be decomposed into a sum of representations in $\hat{G}$.

Example 6.1 Fix a group G and a representation ρ. Let $\rho_1, \ldots, \rho_k$ be the irreducible representations of G. Desiring to know how ρ decomposes in these ρ_i, we compute

$$n_i = \langle \chi_\rho, \chi_i \rangle$$

for each $i = 1, \ldots, k$. Then $\rho = n_1\rho_1 \oplus \cdots \oplus n_k\rho_k$, and after a unitary change of basis, the diagonal of the matrix $\rho(g)$ consists of n_1 copies of $\rho_1(g)$, followed by n_2 copies of $\rho_2(g)$, and so on. ■

There are two representations that every group has:

The Trivial Representation. The trivial representation $\mathbf{1}_G$ maps every group element $g \in G$ to the 1 by 1 matrix (1). One feature of the trivial representation is that

$\sum_g \mathbf{1}_G(g)$ is the 1×1 matrix $(|G|)$; this sum is the zero matrix for any other irreducible representation.

The Regular Representation. We take a vector space V with a basis vector e_g for every element $g \in G$. The regular representation $\mathrm{reg}_G : G \to \mathrm{GL}(V)$ is defined by $\mathrm{reg}_G(g) : e_x \mapsto e_{gx}$ for any $x \in G$. It has dimension $|G|$. With the basis above, for any $g \in G$, $\mathrm{reg}_G(g)$ is a permutation matrix.

An important fact about the regular representation is that it contains every irreducible representation of G. In particular, if $\rho_1, \ldots, \rho_k$ are the irreducible representations of G with dimensions $d_1, \ldots, d_k$, then

$$\rho_{\mathrm{reg}} = d_1\rho_1 \oplus \cdots \oplus d_k\rho_k,$$

that is, the regular representation contains each irreducible representation ρ_i exactly d_i times. Counting dimensions,

$$|G| = \sum_i d_i^2. \tag{6.1}$$

The main tool in quantum polynomial time algorithms is the Fourier transform.

Definition 6.1 *Let $f : G \to \mathbb{C}$. The* Fourier transform of f at the irreducible representation ρ *is the $d_\rho \times d_\rho$ matrix*

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{|G|}} \sum_{g \in G} f(g)\rho(g).$$

We refer to the collection of matrices $\langle \hat{f}(\rho) \rangle_{\rho \in \hat{G}}$ as the *Fourier transform* of f. Thus f is mapped into $|\hat{G}|$ matrices of different dimensions. The total number of entries in these matrices is $\sum d_\rho^2 = |G|$, by equation 6.1 above. The Fourier transform is linear in f;

with the constant used above (i.e. $\sqrt{d_\rho/|G|}$) it is in fact unitary, taking the $|G|$ complex numbers $\langle f(g)\rangle_{g\in G}$ to $|G|$ complex numbers organized into matrices.

A familiar case in computer science is when the group is cyclic of order n. Then the linear transformation (i.e. the Fourier transform) is a Vandermonde matrix with n-th roots of unity and the matrices are 1-by-1.

In the quantum setting we identify the superposition $\sum_{g\in G} f_g|g\rangle$ with the function $f : G \to \mathbb{C}$ defined by $f(g) = f_g$. In this notation, $\sum_{g\in G} f(g)|g\rangle$ is mapped under the Fourier transform to $\sum_{\rho\in\hat{G},1\le i,j\le d_\rho} \hat{f}(\rho)_{i,j}|\rho,i,j\rangle$. We remind the reader that $\hat{f}(\rho)_{i,j}$ is a complex number. When the first portion of this triple is measured, we observe $\rho \in \hat{G}$ with probability

$$\sum_{1\le i,j\le d_\rho} \left|\hat{f}(\rho)_{i,j}\right|^2 = \left\|\hat{f}(\rho)\right\|^2$$

where $\|A\|$ is the natural norm given by $\|A\|^2 = \mathrm{Tr}A^*A$.

Let f be the indicator function of a left coset of H in G, i.e. for some $c \in G$,

$$f(g) = \begin{cases} \frac{1}{\sqrt{|H|}} & \text{if } g \in cH \\ 0 & \text{otherwise.} \end{cases}$$

Our goal is to understand the Fourier transform of f. As mentioned above, the probability of observing ρ is $\left\|\hat{f}(\rho)\right\|^2 = \sum_{i,j} |(\hat{f}(\rho))_{i,j}|^2$. Our choice to measure only the representation ρ (and not the matrix indices) depends on the following key fact about the Fourier transform, also relevant to the abelian solution:

Claim 6.1 *The probability of observing ρ is independent of the coset.*

Proof: $\hat{f}(\rho) = \sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \sum_{h \in H} \rho(ch) = \rho(c)\sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \sum_{h \in H} \rho(h)$ and, since $\rho(c)$ is a unitary matrix,

$$\left\| \hat{f}(\rho) \right\|^2 = \left\| \rho(c)\sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \sum_{h \in H} \rho(h) \right\|^2 = \left\| \sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \sum_{h \in H} \rho(h) \right\|^2 .$$

■

Given this, we may assume that our function f is positive on the subgroup H itself, and zero elsewhere.

6.2 The Probability Distribution over Representations

The primary question is that of the probability of observing ρ. We have seen that this is determined by $\sum_{h \in H} \rho(h)$ which, being a sum of the linear transformations $\rho(h)$, is a linear transformation. We begin by showing that it is a projection:

Claim 6.2 *$\hat{f}(\rho)$ is a projection.*

It is actually a scalar multiple of a projection, but we will ignore factors for now. With the right basis, then, $\hat{f}(\rho)$ will be diagonal, and the diagonal entries will consist of ones and zeros. The probability of observing a particular representation ρ will then correspond to the number of ones appearing on the diagonal (i.e., on the dimension of the image of $\hat{f}(\rho)$).

Given an irreducible representation ρ of G, we are interested in the sum of the matrices $\rho(h)$ for all $h \in H$. Since we only evaluate ρ on H, we can instead consider $\mathrm{Res}_H \rho$ without changing anything. As mentioned before, though ρ is irreducible (over G), $\mathrm{Res}_H \rho$ may not be irreducible on H. We may, however, decompose $\mathrm{Res}_H \rho$ into irreducible

representations over H. Then the Fourier transform of f at ρ is comprised of blocks, each corresponding to a representation in the decomposition of $\mathrm{Res}_H \rho$. In particular, the matrix $\sum_{h \in H} \rho(h)$ is:

$$U \begin{bmatrix} \sum_{h \in H} \sigma_1(h) & 0 & \cdots & 0 \\ 0 & \sum_{h \in H} \sigma_2(h) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{h \in H} \sigma_t(h) \end{bmatrix} U^* \tag{6.2}$$

for some unitary transformation U and some irreducible representations σ_i of H (with possible repetitions). We know that the sum is nonzero only when the irreducible representation is trivial, in which case it is $|H|$. Then the probability of observing ρ is

$$\begin{aligned} \left\| (\hat{f}(\rho)) \right\|^2 = \left\| \sqrt{\frac{d_\rho}{|G|}} \frac{1}{\sqrt{|H|}} \sum_{h \in H} \rho(h) \right\|^2 &= \frac{d_\rho}{|G|} \frac{1}{|H|} |H|^2 \langle \chi_\rho, \chi_{\mathbf{1}_H} \rangle_H \\ &= \frac{|H|}{|G|} d_\rho \langle \chi_\rho, \chi_{\mathbf{1}_H} \rangle_H. \end{aligned}$$

We have proved:

Theorem 6.1 *For any subgroup $H \leq G$, $\left\| \hat{f}(\rho) \right\|^2 = \frac{|H|}{|G|} d_\rho \langle \chi_\rho, \chi_{\mathbf{1}_H} \rangle_H$.*

Observe that one consequence of the theorem is that the probability of observing a representation ρ depends only on the character of ρ. It turns out that characters are class functions, i.e. $\chi(g) = \chi(hgh^{-1})$ for any character χ and $h, g \in G$. Hence conjugate subgroups (gHg^{-1} for some $g \in G$ is a conjugate subgroup of H) produce exactly the same distribution; this rules out using the paradigm of Algorithm 6.1 with measuring representations alone to solve the HSP for any group containing a subgroup that is not normal.

6.3 A Positive Result: Normal Subgroups

In this section we show that Fourier sampling a polynomial number of coset states suffices to reconstruct normal subgroups. Using Theorem 6.1, which describes the probability of measuring a representation, the main point is that when the subgroup is normal, the Fourier transform at each representation takes a nice form. Namely, it is basically either the identity or zero, depending on whether or not H is in the kernel[1] of ρ:

Lemma 6.1 *Let $H \trianglelefteq G$. Then* $\hat{f}(\rho) = \begin{cases} c \cdot I & \text{if } H \subseteq \ker \rho, \text{ where } c = \frac{|H|}{|G|} d_\rho. \\ 0 \text{ matrix} & \text{otherwise} \end{cases}$

Proof: The lemma follows from a simple application of Schur's lemma [47]. We will give the whole proof. Two alternative proofs are in [33]. If $H \subseteq \ker \rho$ then the value of c follows from the discussion in the previous section.

Suppose $H \not\subseteq \ker \rho$. We will show that $\hat{f}(\rho)$ must be the zero map. Let $V = W_1 \bigoplus W_2$, where H fixes W_1 and kills W_2. Since ρ is irreducible over G, we can choose a vector $w_1' \in W_1$ and $g \in G$ such that $\rho_g w_1' = w_1 + w_2$, with $w_i \in W_i$, and $w_2 \neq 0$. Since H fixes W_1 we have

$$w_1 + w_2 = \rho_g w_1' = \rho_g \frac{1}{|H|} \sum_{h \in H} \rho_h w_1' = \frac{1}{|H|} \sum_{h \in H} \rho_h \rho_g w_1' = \frac{1}{|H|} \sum_{h \in H} \rho_h (w_1 + w_2) = w_1,$$

since H is normal in G. This is a contradiction unless $\sum_{h \in H} \rho_h$ is the zero map. ∎

We need one more fact to prove the theorem, which is that each time we sample we have a good chance of getting closer to H. Let $H_i = \bigcap_{j=1}^{i} \ker \rho_j$, the normal subgroup defined by the first i samples.

[1] We remind the reader that a representation σ is a homomorphism $\sigma : G \to GL(V)$. The *kernel* of ρ is the set $\ker(\sigma) = \{g \in G | \sigma(g) = \mathbf{1}_V\}$ and is a normal subgroup of G, which we write $\ker \sigma \trianglelefteq G$.

Claim 6.3 *If $H_i \neq H$ then* $\Pr(H_{i+1} = H_i) \leq \frac{1}{2}$,

Proof: By Lemma 6.1, Theorem 6.1 and Equation 6.1 we have:

$$\Pr(H_i \subseteq \ker \rho_{i+1}) = \sum_{\substack{\rho \in \hat{G} \\ H_i \subseteq \ker \rho}} \Pr(\rho) = \sum_{\substack{\rho \in \hat{G} \\ H_i \subseteq \ker \rho}} \frac{|H|}{|G|} d_\rho^2 = \frac{|H|}{|G|} \sum_{\rho \in \widehat{G/H_i}} d_\rho^2 = \frac{|H|}{|G|} \frac{|G|}{|H_i|} \leq \frac{1}{2}$$

where changing the sum follows from the fact that representations of G that map H_i to the identity can be identified with representations of G/H_i. ■

We now apply a Martingale Bound [40] to prove the theorem.

Theorem 6.2 *Let $\rho_1, \ldots, \rho_k$ be independent random variables distributed according to $\mathcal{D}$ with $k = 4\log_2 |G|$. Then*

$$\Pr[H \neq \bigcap_i \ker \rho_i] \leq 2e^{-\log_2 |G|/8}.$$

Proof: For each $i \in \{1, \ldots, k\}$, let X_i be the indicator random variable taking value 1 if $H_i = H$ or $H_{i+1} \neq H_i$, and zero otherwise. The random variables $X_1, \ldots, X_k$ are not necessarily independent, but by the previous claim, $\Pr[X_i = 0 \,|\, X_1, \ldots, X_{i-1}] \leq \frac{1}{2}$, and we can use a martingale bound. The function "sum" satisfies the Lipschitz condition, so we can apply Azuma's Inequality to conclude that $\sum_i X_i$ does not deviate much from its expected value, which is at least $\frac{k}{2}$. In particular, we have $\Pr\left[|\sum_i X_i - \frac{k}{2}| \geq \lambda\right] \leq 2e^{-\lambda^2/2k}$, so with $\lambda = \log_2(|G|)$ we have $\Pr\left[\sum_{i=0}^{k-1} X_i \leq \log_2(|G|)\right] \leq 2e^{-\log_2(|G|)/8}$. ■

It is also easy to see how the algorithm works when H is not normal.

Theorem 6.3 *For any subgroup H of G, Algorithm 6.1 returns the largest subgroup of H that is normal in G.*

Proof: This proof is due to Vazirani [50]. Let N be the intersection of the kernels so far. Let r_ρ be the rank of $\hat{f}(\rho)$, i.e., $r_\rho = \langle \chi_\rho, \chi_{tr} \rangle$. When $N \not\subseteq H$, we will show that the probability of N being contained in the kernel of the next representation we measure is at most $1/2$, by showing that $\sum_{\rho: N \subseteq \ker \rho} \frac{|H|}{|G|} d_\rho r_\rho \leq \frac{|N \cap H|}{|N|}$, which is at most $1/2$ when $N \not\subseteq H$. If the hidden subgroup had been HN, Theorem 6.1 would imply $\sum_{\rho: \rho \in \hat{G}} \frac{|HN|}{|N|} d_\rho r'_\rho = 1$, where r'_ρ is the number of times the trivial representation of HN appears in ρ. Note that $r'_\rho = r_\rho$ when $N \subseteq \ker \rho$, since $|H \cap N| \sum_{l \in HN} \rho(l) = (\sum_{h \in H} \rho(h))(\sum_{n \in N} \rho(n))$, and $\rho(n)$ is the identity. Since $|HN| \cdot |H \cap N| = |H| \cdot |N|$, we have that $\sum_{\rho: N \subseteq \ker \rho} \frac{|H|}{|G|} d_\rho r'_\rho \leq \sum_{\rho \in \hat{G}} \frac{|H|}{|G|} d_\rho r'_\rho \leq \frac{|H \cap N|}{|N|}$, as desired.

Once $N \subseteq H$, the process stops. First, there is a unique largest subgroup H' of H that is normal in G since if there are two, their product is in H and is normal. Second, we measure representations ρ such that $H' \subseteq \ker \rho$. Let C be a set of coset representatives for H' in H. We have $\sum_{h \in H} \rho(h) = (\sum_{c \in C} \rho(c))(\sum_{h \in H'} \rho(h))$, so by Lemma 6.1 we only see ρ if $\sum_{h \in H'} \rho(h)$ is a multiple of the identity, i.e. only if $H' \subseteq \ker \rho$. ■

6.4 A Negative Result: Determining Triviality in the Symmetric Group

In this section we analyze the HSP algorithm when the group is the symmetric group S_n. We show that Fourier sampling a polynomial number of coset states is insufficient to distinguish the trivial subgroup from an order two subgroup. This implies that this algorithm cannot solve graph isomorphism.

Example 6.2 Reducing graph isomorphism to the HSP over the symmetric group [37]

Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ on vertex sets of size $n/2$, let $G = G_1 \times G_2$ where $V_1 = \{1, \ldots, n/2\}$ and $V_2 = \{n/2 + 1, \ldots, n\}$. Without loss of generality, assume the graphs are connected (otherwise looking at the complement of the graph suffices). G_1 and G_2 are isomorphic iff there exists a map $\phi : V_1 \to V_2$ such that $(i, j) \in E_1$ iff $(\phi(i), \phi(j)) \in E_2$.

Now define a function $f : S_n \to T$ for a set T by $f(\pi) = \pi G$. f is an instance of the HSP. The hidden subgroup H is $\mathrm{Aut}(G)$, the set of isomorphisms from G to G, called the automorphisms of the graph. We can write S_n as a sum of left cosets: $S_n = \sum_i \sigma_i \mathrm{Aut}(G)$. Then for any $\pi \in S_n$, we have $\pi = \sigma_i \phi$ for some i and $\phi \in H$. π_1 and π_2 are in the same coset iff they have the same set of edge violations. ■

We will now show that a special case of this problem cannot even be solved by Fourier sampling coset states. Consider the case when G_1 and G_2 are both rigid, that is, G_1 and G_2 have trivial automorphism groups. If G_1 and G_2 are not isomorphic, $\mathrm{Aut}(G) = \{e\}$, and if they are isomorphic, $\mathrm{Aut}(G) = \{e, \tau\}$ where τ swaps the vertices of G_1 and G_2 in some way, and so τ is a product of $n/2$ disjoint 2-cycles. Let $\mathcal{D}_N$ and $\mathcal{D}_I$ denote the two distributions induced by these two cases. It is enough to decide whether or not $\mathrm{Aut}(G)$ is trivial or nontrivial. We show that even for this particular case of graph isomorphism (and Graph Automorphism) the algorithm fails.

Theorem 6.4 $|\mathcal{D}_I - \mathcal{D}_N|_1 \leq 2^{-\Omega(n)}$.

Proof: We present the proof from [28]. An alternative proof appears in [33]. When $G_1 \not\approx G_2$, $H = \{e\}$, so $\mathcal{D}_N(\rho) = \frac{d_\rho^2}{n!}$ by Theorem 6.1. When $G_1 \approx G_2$, and G_1 and G_2 are both connected and rigid, $H = \{e, \tau\}$. By Theorem 6.1

$$\mathcal{D}_I(\rho) \;=\; \frac{|H|}{|G|} d_\rho \left\langle \chi_{\mathbf{1}}, \chi_\rho \right\rangle_H$$

H has only two elements, e and τ, hence

$$\left\langle \chi_{\mathbf{1}}, \chi_\rho \right\rangle_H = \frac{1}{2}(\chi_\rho(e) + \chi_\rho(\tau)) = \frac{1}{2}(d_\rho + \chi_\rho(\tau)).$$

That is, $\mathcal{D}_I(\rho) = \frac{d_\rho}{n!}(d_\rho + \chi_\rho(\tau))$ and so,

$$\sum_\rho |\mathcal{D}_I(\rho) - \mathcal{D}_N(\rho)| \;=\; \frac{1}{n!}\sum_\rho d_\rho |\chi_\rho(\tau)| \le \frac{1}{n!}\sqrt{\sum_\rho d_\rho^2}\sqrt{\sum_\rho |\chi_\rho(\tau)|^2} \le \frac{1}{\sqrt{n!}}\sqrt{\sum_\rho |\chi_\rho(\tau)|^2}$$

by the Cauchy-Schwartz Inequality and Equation 6.1.

To bound $\sum_\rho |\chi_\rho(\tau)|^2$ we use the fact that the character table with the right normalization factors is a unitary matrix indexed by conjugacy classes and representations. In particular, the row for the conjugacy class of τ, $\{\tau\}$, has normalization factor $\sqrt{|\{\tau\}|/|S_n|}$. The inner product of this row with itself is one, so we have:

$$\frac{1}{\sqrt{n!}}\sqrt{\sum_\rho |\chi_\rho(\tau)|^2} \;=\; \frac{1}{\sqrt{|\{\tau\}|}} = \sqrt{\frac{|C_{S_n}(\tau)|}{n!}} = \sqrt{\frac{2^{(n/2)}(n/2)!}{n!}} \le 2^{-\Omega(n)},$$

where $C_{S_n}(\tau) = \{\sigma \in S_n | \tau\sigma = \sigma\tau\}$ is the centralizer of τ. ■