Finally, an application of Proposition 4.1 allows us to remove the condition of n being convenient from the main term on the right hand side above. This completes the proof of Proposition 4.3, up to the proof of Proposition 4.4, which we take up in the next section. □

5. Counting solutions to congruences: Generalization of Proposition 4.4

We devote this section to establishing a general version of Proposition 4.4 which shall also be useful while dealing with the contribution of the inconvenient n in the proof of Theorems 2.1 to 2.4. To do this, we shall primarily make use of two bounds on character sums, which we state in the next two propositions. In this section, we deviate from the notation and hypotheses set up in the introduction, assuming *only* what is introduced in the rest of the section.

Proposition 5.1. *Let ℓ be a prime, χ a Dirichlet character mod ℓ, and $F \in \mathbb{Z}[T]$ a nonconstant polynomial which is not congruent mod ℓ to a polynomial of the form $c \cdot G(T)^{\mathrm{ord}(\chi)}$ for some $c \in \mathbb{F}_\ell$ and $G \in \mathbb{F}_\ell[T]$, where $\mathrm{ord}(\chi)$ denotes the order of the character χ. Then*

$$\left| \sum_{u \bmod \ell} \chi(F(u)) \right| \leq (d-1)\sqrt{\ell},$$

where d is the degree of the largest squarefree divisor of F.

This is a version of the Weil bounds and is a special case of [49, Corollary 2.3] (see also [9], [50] and [41] for older results). We will also need an analogue of the above result for character sums to higher prime power moduli, and this input is provided by the following consequences of Theorems 1.2 and 7.1 and eqn. (1.15) in work of Cochrane [6] (see [7] for related results).

In what follows, for a polynomial $H \in \mathbb{Z}[T]$, we denote by H' or $H'(T)$ the formal derivative of H. Given a prime ℓ, by the ℓ-critical polynomial associated to H we shall mean the polynomial $\mathcal{C}_H := \ell^{-\mathrm{ord}_\ell(H')} H'$, which has integer coefficients and can be considered as a nonzero element of the ring $\mathbb{F}_\ell[T]$. Moreover, if H is not identically zero in $\mathbb{F}_\ell[T]$ (i.e., if $\mathrm{ord}_\ell(H) = 0$), then by the ℓ-critical points of H, we shall mean the set $\mathcal{A}(H;\ell) \subset \mathbb{F}_\ell$ of zeros of the polynomial $\mathcal{C}_H$ which are not zeros of H (both polynomials considered mod ℓ). Finally, for any $\theta \in \mathbb{F}_\ell$, we use $\mu_\theta(H)$ to denote the multiplicity of θ as a zero of H.

Proposition 5.2. *Let ℓ be a prime, $g \in \mathbb{Z}[T]$ a nonconstant polynomial, and $t := \mathrm{ord}_\ell(g')$. Consider an integer $e \geq t+2$ and a primitive character χ mod ℓ^e. Let $M := \max_{\theta \in \mathcal{A}(g;\ell)} \mu_\theta(\mathcal{C}_g)$ be the maximum multiplicity of an ℓ-critical point.*

(i) *For odd ℓ, we have $|\sum_{u \bmod \ell^e} \chi(g(u))| \leq \left(\sum_{\theta \in \mathcal{A}(g;\ell)} \mu_\theta(\mathcal{C}_g)\right) \ell^{t/(M+1)}\ \ell^{e(1-1/(M+1))}$.*

(ii) *For $\ell = 2$ and $e \geq t+3$, we have $|\sum_{u \bmod 2^e} \chi(g(u))| \leq (12.5) 2^{t/(M+1)}\ 2^{e(1-1/(M+1))}$. In fact, the sum is zero if g has no 2-critical points.*

In order to make use of the aforementioned bounds, we will need to understand the quantities that appear when we apply them. The following observations enable us to do this.

Proposition 5.3. *Let $\{F_i\}_{i=1}^K \subset \mathbb{Z}[T]$ be nonconstant and multiplicatively independent. There exists a constant $C_1 := C_1(F_1, \ldots, F_K) \in \mathbb{N}$ such that all of the following hold:*

(a) *For any prime ℓ, there are $O(1)$ many tuples $(A_1,\dots,A_K)\in[\ell-1]^K$ for which $F_1^{A_1}\cdots F_K^{A_K}$ is of the form $c\cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$ for some $c\in\mathbb{F}_\ell$ and $G\in\mathbb{F}_\ell[T]$; here, the implied constant depends at most on $\{F_i\}_{i=1}^K$. In fact, if $\ell>C_1$ and $\gcd(\ell-1,\beta(F_1,\dots,F_K))=1$, then the only such tuple is $(A_1,\dots,A_K)=(\ell-1,\dots,\ell-1)$.*

(b) *For any prime ℓ and any $(A_1,\dots,A_K)\in\mathbb{N}^K$ satisfying $\mathrm{ord}_\ell(\prod_{i=1}^K F_i)=0$ and $(A_1,\dots,A_K)\not\equiv(0,\dots,0)\pmod{\ell}$, we have in the two cases below,*

$$\tau(\ell)\coloneqq\mathrm{ord}_\ell\left((T^{\varphi(\ell^r)}F_1(T)^{A_1}\cdots F_K(T)^{A_K})'\right)=\mathrm{ord}_\ell(\widetilde{F}(T))\begin{cases}=0, & \text{if } \ell>C_1, r\geq 2\\ \leq C_1, & \text{if } \ell\leq C_1, \mathrm{ord}_\ell\left(\prod_{i=1}^K F_i\right)=0, r\geq C_1+2,\end{cases}\tag{5.1}$$

where $\widetilde{F}(T)\coloneqq\sum_{i=1}^K A_iF_i'(T)\prod_{\substack{1\leq j\leq K\\ j\neq i}}F_j(T)$. In either of the two cases above, any root $\theta\in\mathbb{F}_\ell$ of the polynomial $\mathcal{C}_\ell(T)\coloneqq\ell^{-\tau(\ell)}(T^{\varphi(\ell^r)}F_1(T)^{A_1}\cdots F_K(T)^{A_K})'$ which is not a root of $T\prod_{i=1}^K F_i(T)$, must be a root of the polynomial $\ell^{-\tau(\ell)}\widetilde{F}(T)$ of the same multiplicity.[5]

Proof. We start by writing $F_i=:r_i\prod_{j=1}^M G_j^{\mu_{ij}}$ as in the introduction, so that $r_i\in\mathbb{Z}$ and $G_1,\dots,G_M\in\mathbb{Z}[T]$ are irreducible, primitive and pairwise coprime, and $M=\omega(F_1\cdots F_K)$. Recall that $M\geq K$ and that the exponent matrix $E_0(F_1,\dots,F_K)$ has $\mathbb{Q}$-linearly independent columns, making $\beta(F_1,\dots,F_K)$ a nonzero integer. Further, since G_j are pairwise coprime irreducibles, the resultants $\mathrm{Res}(G_j,G_{j'})$ and discriminants $\mathrm{disc}(G_j)$ are nonzero integers for all $j\neq j'\in[M]$. Note that for any prime ℓ not dividing the leading coefficient of any G_j and not dividing $\prod_{1\leq j\leq M}\mathrm{disc}(G_j)\cdot\prod_{1\leq j\neq j'\leq M}\mathrm{Res}(G_j,G_{j'})$, the product $\prod_{j=1}^M G_j$ is separable in $\mathbb{F}_\ell[T]$.

We also observe that since $(F_1^{c_1}\cdots F_K^{c_K})'=\left(\prod_{i=1}^K F_i^{c_i-1}\right)\sum_{i=1}^K c_iF_i'\prod_{\substack{1\leq j\leq K\\ j\neq i}}F_j$, the multiplicative independence of the polynomials $\{F_i\}_{i=1}^K$ forces the polynomials $\left\{F_i'\prod_{\substack{1\leq j\leq K\\ j\neq i}}F_j\right\}_{i=1}^K\subset\mathbb{Z}[T]$ to be $\mathbb{Q}$-linearly independent. Writing $D\coloneqq\deg(F_1\cdots F_K)$ and $F_i'(T)\prod_{\substack{1\leq j\leq K\\ j\neq i}}F_j(T)=:\sum_{j=0}^{D-1}c_{i,j}T^j$ for some $\{c_{i,j}\}_{0\leq j\leq D-1}\subset\mathbb{Z}$, we find that the columns of the matrix

$$M_1\coloneqq M_1(F_1,\dots,F_K)\coloneqq\begin{pmatrix}c_{1,0} & \cdots & c_{K,0}\\ \cdots & \cdots & \cdots\\ \cdots & \cdots & \cdots\\ c_{1,D-1} & \cdots & c_{K,D-1}\end{pmatrix}\in\mathbb{M}_{D\times K}(\mathbb{Z})\tag{5.2}$$

must be $\mathbb{Q}$-linearly independent. Consequently, the last diagonal entry $\widetilde{\beta}\coloneqq\widetilde{\beta}(F_1,\dots,F_K)\in\mathbb{Z}\setminus\{0\}$ is the largest invariant factor of M_1 (in size).

We now let $C_1\coloneqq C_1(F_1,\dots,F_K)$ be any positive integer exceeding $\max\{|\widetilde{\beta}|,2\}$ such that for any $\ell>C_1$, ℓ divides neither the product $\prod_{j=1}^M\mathrm{disc}(G_j)\cdot\prod_{1\leq j\neq j'\leq M}\mathrm{Res}(G_j,G_{j'})\in\mathbb{Z}\setminus\{0\}$ nor the leading coefficient of any of $F_1,\dots,F_K$ (hence also none of the leading coefficients of $G_1,\dots,G_M$), and we have $\mathrm{ord}_\ell(F_1\cdots F_K)=0$. We claim that any such C_1 satisfies the properties in the statement of the proposition.

[5]Once again, the last three polynomials are being considered as nonzero elements of $\mathbb{F}_\ell[T]$.

Proof of (a). We may assume that $\ell > C_1$. Let $\beta := \beta(F_1, \dots, F_K)$. As mentioned before, the conditions defining C_1 force $G_1, \dots, G_M$ to be pairwise coprime in $\mathbb{F}_\ell[T]$. Let $(A_1, \dots, A_K) \neq (0, \dots, 0)$ be any tuple of nonnegative integers for which $F_1^{A_1} \cdots F_K^{A_K}$ is of the form $c \cdot G^{\ell-1}$ in $\mathbb{F}_\ell[T]$ for some $c \in \mathbb{F}_\ell$ and $G \in \mathbb{F}_\ell[T]$. We claim that $A_1, \dots, A_K$ must all be divisible by $(\ell-1)/d_1$ where $d_1 := \gcd(\ell-1, \beta)$. This will be enough to complete the proof of (a), since there are no more than $d_1^K \leq |\beta|^K \ll 1$ many tuples $(A_1, \dots, A_K) \in [\ell-1]^K$ satisfying this latter property.

To establish the above claim, we may assume without loss of generality that G is monic, and note that $c \in \mathbb{F}_\ell^\times$ since $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$ by definition of C_1. Write each G_j as $\lambda_j H_j$ in the ring $\mathbb{F}_\ell[T]$, for some $\lambda_j \in \mathbb{F}_\ell^\times$ and nonconstant monic $H_j \in \mathbb{F}_\ell[T]$ (which can be done since ℓ doesn't divide the leading coefficient of any G_j). Then $F_i = r_i \prod_{j=1}^M G_j^{\mu_{ij}} = \rho_i \prod_{j=1}^M H_j^{\mu_{ij}}$ for some $\rho_i \in \mathbb{F}_\ell^\times$. Since $c \cdot G^{\ell-1} = \prod_{i=1}^K F_i^{A_i} = \left(\prod_{i=1}^K \rho_i^{A_i}\right) \cdot \prod_{1 \leq j \leq M} H_j^{\sum_{i=1}^K \mu_{ij} A_i}$ in $\mathbb{F}_\ell[T]$, and $G, H_1, \dots, H_M$ are all monic, we find that $G^{\ell-1} = \prod_{1 \leq j \leq M} H_j^{\sum_{i=1}^K \mu_{ij} A_i}$. But now since $\prod_{1 \leq j \leq M} G_j$ is separable in $\mathbb{F}_\ell[T]$, so is $\prod_{1 \leq j \leq M} H_j$, and we deduce that $\sum_{i=1}^K \mu_{ij} A_i \equiv 0 \pmod{\ell-1}$ for each $j \in [M]$. This can be rewritten as the matrix congruence $(0 \cdot \cdots \cdot 0)^\top \equiv E_0 (A_1 \cdots A_K)^\top \pmod{\ell-1}$, where each side is an $M \times 1$ matrix and $Y^\top$ denotes the transpose of a matrix Y.

Now since $M \geq K$ and E_0 has full rank, there exist $P_0 \in GL_{M \times M}(\mathbb{Z})$ and $R_0 \in GL_{K \times K}(\mathbb{Z})$ for which $P_0 E_0 R_0$ is the Smith Normal Form $\mathrm{diag}(\beta_1, \dots, \beta_K)$ of E_0, with $\beta_1, \dots, \beta_K \in \mathbb{Z} \setminus \{0\}$ being the invariant factors of E_0, so that $\beta_i \mid \beta_{i+1}$ for all $1 \leq i < K$ and $\beta = \beta(F_1, \dots, F_K) = \beta_K$. Thus $P_0 E_0 = \mathrm{diag}(\beta_1, \dots, \beta_K) R_0^{-1}$ and writing $(q_{ij})_{1 \leq i,j \leq K} := R_0^{-1}$, we find that

$$\begin{pmatrix} 0 \\ \cdots \\ \cdots \\ 0 \end{pmatrix}_{M \times 1} \equiv P_0 E_0 \begin{pmatrix} A_1 \\ \cdots \\ A_K \end{pmatrix}_{K \times 1} \equiv \begin{pmatrix} \beta_1(q_{11}A_1 + \cdots + q_{1K}A_K) \\ \cdots \\ \beta_K(q_{K1}A_1 + \cdots + q_{KK}A_K) \\ 0 \\ \cdots \\ 0 \end{pmatrix}_{M \times 1} \pmod{\ell-1}.$$

Hence for each $i \in [K]$, $\beta_i(q_{i1}A_1 + \cdots + q_{iK}A_K) \equiv 0 \pmod{\ell-1}$, so that $(\ell-1)/\gcd(\ell-1, \beta_i)$ divides $q_{i1}A_1 + \cdots + q_{iK}A_K$. But since $\beta_i \mid \beta_K$, it follows that $(\ell-1)/\gcd(\ell-1, \beta_K) = (\ell-1)/d_1$ also divides $q_{i1}A_1 + \cdots + q_{iK}A_K$ for each $i \in [K]$. We obtain

$$\begin{pmatrix} 0 \\ \cdots \\ 0 \end{pmatrix}_{K \times 1} \equiv \begin{pmatrix} q_{11}A_1 + \cdots + q_{1K}A_K \\ \cdots \\ q_{K1}A_1 + \cdots + q_{KK}A_K \end{pmatrix}_{K \times 1} \equiv R_0^{-1} \begin{pmatrix} A_1 \\ \cdots \\ A_K \end{pmatrix}_{K \times 1} \left(\mathrm{mod}\ \ \frac{\ell-1}{d_1}\right), \tag{5.3}$$

establishing the desired claim that $(A_1, \dots, A_K) \equiv (0, \dots, 0) \left(\mathrm{mod}\ \ \frac{\ell-1}{d_1}\right)$.

Proof of (b). We start by noting that

$$(T^{\varphi(\ell^r)} F_1(T)^{A_1} \cdots F_K(T)^{A_K})' = \varphi(\ell^r) T^{\varphi(\ell^r)-1} \prod_{i=1}^K F_i(T)^{A_i} + T^{\varphi(\ell^r)} \left(\prod_{i=1}^K F_i(T)^{A_i - 1}\right) \widetilde{F}(T), \tag{5.4}$$

where $\widetilde{F}(T)$ is as in the statement of the proposition. We claim that $\mathrm{ord}_\ell(\widetilde{F}) \leq \mathbb{1}_{\ell \leq C_1} C_1$ for all primes ℓ satisfying $\mathrm{ord}_\ell(F_1 \cdots F_K) = 0$ and for all nonnegative integers $A_1, \dots, A_K$