L-FUNCTIONS OF EXPONENTIAL SUMS OVER ONE-DIMENSIONAL AFFINOIDS: NEWTON OVER HODGE

HUI JUNE ZHU

ABSTRACT. This paper proves a sharp lower bound for Newton polygons of L-functions of exponential sums of one-variable rational functions. Let p be a prime and let $\overline{\mathbb{F}}_p$ be the algebraic closure of the finite field of p elements. Let $\overline{f}(x)$ be any one-variable rational function over $\overline{\mathbb{F}}_p$ with ℓ poles of orders $d_1, \ldots, d_\ell$. Suppose p is coprime to $d_1 \cdots d_\ell$. We prove that there exists a tight lower bound which we call Hodge polygon, depending only on the d_j's, to the Newton polygon of L-function of exponential sums of $\overline{f}(x)$. Moreover, we show that for any $\overline{f}(x)$ these two polygons coincide if and only if $p \equiv 1 \bmod d_j$ for every $1 \leq j \leq \ell$. As a corollary, we obtain a tight lower bound for the p-adic Newton polygon of zeta-function of an Artin-Schreier curve given by affine equations $y^p - y = \overline{f}(x)$.

1. INTRODUCTION

Let $\mathbb{A}$ be the space of rational functions in one variable x with ℓ distinct poles (say at $P_1, P_2, \ldots, P_\ell$) of orders $d_1, \ldots, d_\ell \geq 1$ on the projective line. For any field K, we denote by $\mathbb{A}(K)$ the set of all rational functions of the form $\sum_{j=1}^{\ell} \sum_{i=1}^{d_j} a_{j,i}(x - P_j)^{-i}$, where coefficients $a_{j,i} \in K$, poles $P_j \in K \cup \{\infty\}$ and $\prod_{j=1}^{\ell} a_{j,d_j} \neq 0$ (we set $(x - \infty)^{-i} = x^i$ for the point at ∞). Naturally one may consider $\mathbb{A}$ as a quasi-affine space parameterized by coefficients $a_{j,i}$ for all $i \geq 1, 1 \leq j \leq \ell$ and poles P_j for $1 \leq j \leq \ell$. Let the *Hodge polygon* of $\mathbb{A}$, denoted by $\mathrm{HP}(\mathbb{A})$, be the end-to-end join of line segments of horizontal length 1 with slopes listed below:

$$\overbrace{0,\ldots,0}^{\ell-1};\overbrace{1,\ldots,1}^{\ell-1};\overbrace{\frac{1}{d_1},\cdots,\frac{d_1-1}{d_1}}^{d_1-1};\overbrace{\frac{1}{d_2},\cdots,\frac{d_2-1}{d_2}}^{d_2-1};\ldots\ldots;\overbrace{\frac{1}{d_\ell},\cdots,\frac{d_\ell-1}{d_\ell}}^{d_\ell-1}. \tag{1}$$

They are joined in a nondecreasing order from left to right starting from the origin on $\mathbb{R}^2$. Let $d := \sum_{j=1}^{\ell} d_j + \ell - 2$. So $\mathrm{HP}(\mathbb{A})$ is a lower convex hull in $\mathbb{R}^2$ with endpoints $(0,0)$ and $(d, d/2)$.

For any prime p let $E(x) = \exp(\sum_{i=0}^{\infty} \frac{x^{p^i}}{p^i})$ be the p-adic Artin-Hasse exponential function. Let γ be a p-adic root of $\log(E(x))$ with $\mathrm{ord}_p\gamma = \frac{1}{p-1}$. Then $E(\gamma)$ is a primitive p-th root of unity. We fix it and denote it by ζ_p.

Date: Jan 16, 2004.

2000 *Mathematics Subject Classification.* 11,14.

Key words and phrases. Newton polygon, Hodge polygon, L-function, exponential sums, zeta function of Artin-Schreier curves, Dwork trace formula, Monsky-Reich trace formula.

In this paper we let p be a prime coprime to $\prod_{i=1}^{\ell} d_i$ and let a be a positive integer. Let $q=p^a$. Let

$$\overline{f}(x) \quad := \quad \sum_{j=1}^{\ell}\sum_{i=1}^{d_j} \overline{a}_{j,i}(x-\overline{P}_j)^{-i}, \tag{2}$$

where $\overline{a}_{j,i}\in\mathbb{F}_q, \overline{P}_j\in\mathbb{F}_q\cup\{\infty\}$ for every i,j. Let $\overline{g}(x):=\prod_{\overline{P}_j\neq\infty}(x-\overline{P}_j)\in\mathbb{F}_q[x]$. For any positive integer k, the k-th exponential sum of $\overline{f}(x)\in\mathbb{F}_q(x)$ is $S_k(\overline{f}) := \sum \zeta_p^{\mathrm{Tr}_{\mathbb{F}_{q^k}/\mathbb{F}_p}(\overline{f}(x))}$ where the sum ranges over all x in $\mathbb{F}_{q^k}$ such that $\overline{g}(x)\neq 0$. The L-function of the exponential sum of $\overline{f}$ is defined by

$$L(\overline{f};T) := \exp(\sum_{k=1}^{\infty} S_k(\overline{f})\frac{T^k}{k}). \tag{3}$$

It is well known that the L-function is a polynomial in $\mathbb{Z}[\zeta_p][T]$ of degree d (e.g., by combining the Weil Conjecture for curves with the argument in [29] between Remark 1.2 and Corollary 1.3). One may write

$$L(\overline{f};T) \quad = \quad 1+b_1T+b_2T^2+\ldots+b_dT^d\in\mathbb{Z}[\zeta_p][T]. \tag{4}$$

Define the *Newton polygon* of the L-function of $\overline{f}$ over $\mathbb{F}_q$ as the lower convex hull in $\mathbb{R}^2$ of the points $(n,\mathrm{ord}_q(b_n))$ with $0\leq n\leq d$, where we set $b_0=1$ and $\mathrm{ord}_q(\cdot):=\mathrm{ord}_p(\cdot)/a$. We denote it by $\mathrm{NP}(\overline{f};\mathbb{F}_q)$. One notes immediately that the Newton polygon $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ and the Hodge polygon $\mathrm{HP}(\mathbb{A})$ have the same endpoints $(0,0)$ and $(d,d/2)$. Let $\mathrm{lcm}(d_j)$ denote the least common multiple of d_j's for all $1\leq j\leq\ell$. The main result of the present paper is the following.

Theorem 1.1. *Let notation be as above. For any rational function $\overline{f}\in\mathbb{A}(\mathbb{F}_q)$, the Newton polygon $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ lies over the Hodge polygon $\mathrm{HP}(\mathbb{A})$, and their endpoints meet. Moreover, for any $\overline{f}\in\mathbb{A}(\mathbb{F}_q)$ one has $\mathrm{NP}(\overline{f};\mathbb{F}_q)=\mathrm{HP}(\mathbb{A})$ if and only if $p\equiv 1 \bmod (\mathrm{lcm}\ d_j)$.*

Remark 1.2. The first part (i.e., Newton over Hodge) of Theorem 1.1 was a conjecture of Adolphson-Sperber and Bjorn Poonen, described to the author independently in 2001. The case $\ell=1$ is known (see [25] or [28]). The case $\ell=2$ and $\overline{f}(x)$ has only poles at ∞ and 0 (i.e., $\overline{f}(x)$ is a one variable Laurent polynomial) was obtained first by Robba (see [20, Theorems 7.2 and 7.5]). Theorem 1.1 is an analog of Katz-type conjectures (see [9, Theorem 2.3.1] and [14]).

Below we shall discuss some applications of our result in algebraic geometry. A question that remains open is whether there is a curve in every Newton polygon stratus in the moduli space of curves over $\overline{\mathbb{F}}_p$ (for every p). Recently [26] and [27] gave an affirmative answer to this question for $p=2$ by constructing supersingular curves over $\overline{\mathbb{F}}_2$ via a fibre product of Artin-Schreier curves. It is essential to understand the shape of Newton polygons of Artin-Schreier curves, and in particular, to find a sharp lower bound for them. The Newton polygon of the Artin-Schreier curve $C_{\overline{f}}: y^p-y=\overline{f}(x)$ over $\mathbb{F}_q$ (note that $C_{\overline{f}}$ has genus $d(p-1)/2$) is the normalized p-adic Newton polygon of the numerator of the Zeta function $\mathrm{Zeta}(C_{\overline{f}};T)$ of $C_{\overline{f}}$ (here 'normalized' means taking $\mathrm{ord}_p(\cdot)/a$ as the valuation). Denote this Newton polygon by $\mathrm{NP}(C_{\overline{f}};\mathbb{F}_q)$. Then we have the following corollary.

Corollary 1.3. *Let notation be as in Theorem 1.1. For any $\overline{f} \in \mathbb{A}(\mathbb{F}_q)$ and Artin-Schreier curve $C_{\overline{f}} : y^p - y = \overline{f}(x)$, the Newton polygon $\mathrm{NP}(C_{\overline{f}}; \mathbb{F}_q)$ shrunk by a factor of $1/(p-1)$ (vertically and horizontally) is equal to $\mathrm{NP}(\overline{f}; \mathbb{F}_q)$ and it lies over the Hodge polygon $\mathrm{HP}(\mathbb{A})$. Moreover, for any $\overline{f} \in \mathbb{A}(\mathbb{F}_q)$ the equality holds if and only if $p \equiv 1 \bmod (\mathrm{lcm}\, d_j)$.*

Proof. We shall first give an elementary proof of the following relation between the Zeta function of $C_{\overline{f}}$ and the L-function of $\overline{f}$:

$$\mathrm{Zeta}(C_{\overline{f}}; T) = \frac{\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(L(\overline{f}; T))}{(1-T)(1-qT)} \tag{5}$$

with the norm $\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(\cdot)$ being interpreted as the product of conjugates of the L-function $L(\overline{f}; T)$ in $\mathbb{Q}(\zeta_p)$ over $\mathbb{Q}$, the automorphism acting trivially on the variable T. (One may also see, for example, [3, Section VI, (93)] for some relevant discussion.) Recall that for any integer n one has $\sum_{a \in \mathbb{F}_p} \zeta_p^{an} = p$ or 0 depending on whether n is 0 or not, respectively. For any $k \geq 1$ let $C'_{\overline{f}}$ be the curve $C_{\overline{f}}$ less the ℓ ramification points over $\overline{P}_1 = \infty, \overline{P}_2, \ldots, \overline{P}_\ell$. Write $\mathbb{F}^+_{q^k} := \mathbb{F}_{q^k} - \{\overline{P}_2, \ldots, \overline{P}_\ell\}$. Then

$$\#C'_{\overline{f}}(\mathbb{F}_{q^k}) = \sum_{a \in \mathbb{F}_p} \sum_{x \in \mathbb{F}^+_{q^k}} \zeta_p^{a \mathrm{Tr}(\overline{f}(x))}$$

where $\mathrm{Tr}(\cdot) = \mathrm{Tr}_{\mathbb{F}_{q^k}/\mathbb{F}_p}(\cdot)$. It follows that

$$\begin{aligned}
\mathrm{Zeta}(C_{\overline{f}}; \mathbb{F}_q) &= \exp\left(\sum_{k=1}^{\infty} \left(\ell + \#C'_{\overline{f}}(\mathbb{F}_{q^k})\right) \frac{T^k}{k}\right) \\
&= \exp\left(\sum_{k=1}^{\infty} (1 + q^k + \sum_{a \in \mathbb{F}_p^*} \sum_{x \in \mathbb{F}^+_{q^k}} \zeta_p^{a\mathrm{Tr}(\overline{f}(x))}) \frac{T^k}{k}\right) \\
&= \frac{\prod_{a \in \mathbb{F}_p^*} \left(\exp(\sum_{k=1}^{\infty} (\sum_{x \in \mathbb{F}^+_{q^k}} \zeta_p^{a\mathrm{Tr}(\overline{f}(x))})) \frac{T^k}{k}\right)}{(1-T)(1-qT)} \\
&= \frac{\mathrm{N}_{\mathbb{Q}(\zeta_p)/\mathbb{Q}}(L(\overline{f}; T))}{(1-T)(1-qT)}.
\end{aligned}$$

This proves (5). Our first assertion of the corollary follows from (5) and the rest then follows from the first assertion and Theorem 1.1. □

Remark 1.4. We remark that $\mathrm{NP}(\overline{f}; \mathbb{F}_q)$ and $\mathrm{HP}(\mathbb{A})$ always coincide at the slope-0 segments (of horizontal length $\ell - 1$). Indeed, in the spirit of Corollary 1.3, it suffices to show that the Artin-Schreier curve $C_{\overline{f}} : y^p - y = \overline{f}(x)$ has p-rank equal to $(\ell-1)(p-1)$, which follows from *Deuring-Shafarevic formula* (see, for instance, [5, Corollary 1.8]). By symmetry, their slope-1 segments also coincide.

Now we give an amusing example: By the above, the curve

$$C/\overline{\mathbb{F}}_p : y^p - y = a_{1,2}x^2 + a_{1,1}x + \sum_{j=2}^{\ell} (\frac{a_{j,1}}{(x - \overline{P}_j)^2} + \frac{a_{j,2}}{(x - \overline{P}_j)})$$

for odd prime p and nonzero $a_{j,i}$'s, has its Newton polygon slope 0 (resp. 1) of length $(\ell-1)(p-1)$, and slope $1/2$ of length $\ell(p-1)$.

Finally we comment on our conventions for the proof of this theorem: We first note that for $d_1=\ell=1$ the L-function of f is equal to 1 and we shall exclude this case for the rest of the paper for simplicity; Following conventions often used in algebraic geometry, we set $P_1=\infty$. Recall that one can always move one pole of an Artin-Schreier curve to ∞ by an automorphism of the projective line it covers without altering its zeta function. If $\ell>1$ then we set $P_2=0$ for the rest of the paper. This is not a restriction for our purpose since it is observed by definition that $L(\overline{f}(x+c);T)=L(\overline{f}(x);T)$ for any $c\in\mathbb{F}_q$ so one can always shift $\overline{f}(x)$ so that one of its poles lies at 0. (Note that we then have $\overline{g}(0)=0$.)

This paper is organized as follows. We develop core theory of exponential sums over an affinoid in Section 2. There we determine the size of residue disks and derive an effective trace formula of Dwork-Monsky-Reich. In other words, Section 2 contains fundamentals for the rest of the paper. In Section 3, we present a practical algorithm to estimate the p-adic valuation of our Frobenius matrix. Finally Section 4 is devoted to the proof of Theorem 1.1. In a sequel paper [13] we shall study the asymptotic (as p varies) generic Newton polygons for L-functions of exponential sums of one-variable rational function.

2. Exponential sums over one dimensional affinoids

We generalize Robba's work in [20] from one dimensional annuli to one dimensional affinoids. This differs from Dwork's original approach (see [6, 7]), but is somewhat akin to [1]. Our exposition is (of course) after [15, 20, 21, 16, **?**, 18]. For fundamental material considering non-Archimedean geometry see [4] or [22].

2.1. Preliminaries.

Let $\mathbb{Q}_q$ be the degree a unramified extension of $\mathbb{Q}_p$ and let $\mathbb{Z}_q$ be its ring of integers. Let $\overline{\mathbb{Q}}_p$ be the algebraic closure of $\mathbb{Q}_p$, and let $\overline{\mathbb{Z}}_p$ its ring of integers. Let Ω be the p-adic completion of $\overline{\mathbb{Q}}_p$. Let $\Omega_1=\mathbb{Q}_p(\zeta_p)$ and Ω_a the unique unramified extension of Ω_1 of degree a in Ω. Let $\mathcal{O}_1$ and $\mathcal{O}_a$ be the rings of integers in Ω_1 and Ω_a, respectively. Note that $\mathcal{O}_1=\mathbb{Z}_p[\zeta_p]=\mathbb{Z}_p[\gamma]$. Fix roots γ^{1/d_j} in $\mathcal{O}_a$ for the rest of the paper, and let $\Omega_1':=\Omega_1(\gamma^{1/d_1},\cdots,\gamma^{1/d_\ell})$. Let $\Omega_a':=\Omega_1'\Omega_a$. Let $\mathcal{O}_a'$ and $\mathcal{O}_1'$ be the rings of integers of Ω_a' and Ω_1', respectively. Let $|\cdot|_p$ be the p-adic valuation on Ω_a such that $|p|_p=p^{-1}$. The diagrams below represent these field extensions and the associated extensions of their rings of integers.

$$\begin{tikzcd}
 & & \Omega_a' \arrow[dl, dash] \arrow[dr, dash, "a"] & \\
 & \Omega_a \arrow[dl, dash, "p-1"'] \arrow[dr, dash, "a"] & & \Omega_1' \arrow[dl, dash] \\
\mathbb{Q}_q \arrow[dr, dash, "a"'] & & \Omega_1 \arrow[dl, dash, "p-1"] & \\
 & \mathbb{Q}_p & &
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & \mathcal{O}_a' \arrow[dl, dash] \arrow[dr, dash] & \\
 & \mathcal{O}_a \arrow[dl, dash] \arrow[dr, dash] & & \mathcal{O}_1' \arrow[dl, dash] \\
\mathbb{Z}_q \arrow[dr, dash] & & \mathcal{O}_1 \arrow[dl, dash] & \\
 & \mathbb{Z}_p & &
\end{tikzcd}$$

By taking Teichmüller lifts of coefficients and poles of $\overline{f} \in \mathbb{F}_q[x]$, we get $\hat{f}(x) \in \mathbb{Z}_q[x]$ with $\hat{f}(x) = \sum_{i=1}^{d_1} a_{1,i}x^i + \sum_{j=2}^{\ell}\sum_{i=1}^{d_j} a_{j,i}(x-\hat{P}_j)^{-i}$. Since $\hat{P}_j$ is a Teichmüller lift one has $\hat{P}_j^{\tau} = \hat{P}_j^{p}$. Similarly, let $\hat{g}(x) = \prod_{j=2}^{\ell}(x - \hat{P}_j) \in \mathbb{Z}_q[x]$ be the corresponding Teichmüller lift of $\overline{g}(x) \in \mathbb{F}_q[x]$. Note that $\hat{P}_j^q = \hat{P}_j$ and $\hat{P}_j^q \equiv \overline{P}_j \bmod \mathcal{P}$.

We mainly work on p-adic spaces over Ω_a (or Ω_a'). Let $|\Omega_a|_p$ denote the p-adic value group of Ω_a. Let $\mathbf{P}^1$ be the rigid projective line over Ω_a. For any $\hat{P} \in \Omega_a$ and $r \in |\Omega_a|_p$ let $\mathbf{B}[\hat{P}, r]$ and $\mathbf{B}(\hat{P}, r)$ denote the closed disk and (wide) open disk of radius r about $\hat{P}$ on $\mathbf{P}^1$, that is $\mathbf{B}[\hat{P}, r] := \{X \in \Omega_a | |X - \hat{P}|_p \leq r\}$ and $\mathbf{B}(\hat{P}, r) := \{X \in \Omega_a | |X - \hat{P}|_p < r\}$.

Let $r \in |\Omega_a|_p$ and $0 < r < 1$. For any positive p-power s let $\mathbf{A}_{r,s} := \mathbf{B}[0, 1/r] - \bigcup_{j=2}^{\ell}\mathbf{B}(\hat{P}_j^s, r) = \mathbf{P}^1 - \bigcup_{j=1}^{\ell}\mathbf{B}(\hat{P}_j^s, r)$. So $\mathbf{A}_{r,s} = \{X \in \Omega_a | |X|_p \leq 1/r; |X - \hat{P}_j^s|_p \geq r \text{ for } 2 \leq j \leq \ell, \}$ and it is an affinoid over Ω_a. Let $\mathcal{H}(\cdot)$ be the ring of rigid analytic functions over Ω_a of a given affinoid. Hence, $\mathcal{H}(\mathbf{A}_{r,s})$ is a p-adic Banach space with the natural p-adic supremum norm. For ease of notation, we shall abbreviate $\mathbf{A}_r$ for $\mathbf{A}_{r,1}$ in this paper.

2.2. **The p-adic Mittag-Leffler decomposition.** Let $\mathcal{H}_1(\mathbf{A}_r)$ be the subset of $\mathcal{H}(\mathbf{A}_r)$ consisting of all rigid analytic functions on $\mathbf{B}[0, 1/r]$. For $2 \leq j \leq \ell$ let $\mathcal{H}_j(\mathbf{A}_r)$ be the subset of $\mathcal{H}(\mathbf{A}_r)$ consisting of all rigid analytic functions on $\mathbf{P}^1 - \mathbf{B}(\hat{P}_j, r)$ that are holomorphic at ∞ and vanish at ∞. For any rigid analytic function ξ defined on a subset B of $\mathbf{P}^1$, let $||\xi||_B := \sup_{x\in B}|\xi(x)|_p$ (i.e., supremum norm). This defines a norm on $\mathcal{H}(\mathbf{A}_r)$ and $\mathcal{H}_j(\mathbf{A}_r)$, which are p-adic Banach spaces under the supremum norm $||\xi||_{\mathbf{A}_r}$.

Lemma 2.1 (p-adic Mittag-Leffler)**.** *Let $r \in |\Omega_a|_p$ and $0 < r < 1$. Then the $\mathbf{B}(\hat{P}_j, r)$'s with $1 \leq j \leq \ell$ are mutually disjoint. There is a canonical decomposition of p-adic Banach spaces $\mathcal{H}(\mathbf{A}_r) \cong \bigoplus_{j=1}^{\ell}\mathcal{H}_j(\mathbf{A}_r)$ in the sense that for any $\xi \in \mathcal{H}(\mathbf{A}_r)$ there is a unique $\xi_{\hat{P}_j} \in \mathcal{H}_j(\mathbf{A}_r)$ such that every $\xi - \xi_{\hat{P}_j}$ is analytically expandable to $\mathbf{B}(\hat{P}_j, r)$. Every ξ can be uniquely represented as a sum $\xi = \sum_{j=1}^{\ell}\xi_{\hat{P}_j}$ such that*

$$||\xi||_{\mathbf{A}_r} = \max_{1\leq j\leq \ell}(||\xi_{\hat{P}_j}||_{\mathbf{P}^1-\mathbf{B}(\hat{P}_j,r)}). \tag{6}$$

Proof. We first show the disjointness. Let $j \geq 3$. Since the $\hat{P}_j$'s are Teichmüller lifts in $\mathbb{Z}_q$ with $\hat{P}_j^q = \hat{P}_j$ one has $|\hat{P}_j|_p = 1$. For any $3 \leq i < j \leq \ell$, one first observes easily that $|\hat{P}_i - \hat{P}_j|_p \leq \max(|\hat{P}_i|_p, |\hat{P}_j|_p) = 1$. The hypothesis that $\overline{P}_i \neq \overline{P}_j$ in the residue field of $\mathbb{Z}_q$ implies that $|\hat{P}_i - \hat{P}_j|_p \not< 1$ and hence one has $|\hat{P}_i - \hat{P}_j|_p = 1$.

Let $j \geq 2$. Pick any $\hat{P} \in \mathbf{B}(\hat{P}_j, r)$. If $j = 2$ then $|\hat{P}|_p < r < 1 < 1/r$ so $\hat{P} \in \mathbf{B}[1, 1/r]$; If $j \geq 3$ then $|\hat{P} - \hat{P}_j|_p < r < 1$ and $|\hat{P}_j|_p = 1$ imply that $|\hat{P}|_p = 1 < 1/r$ so $\hat{P} \in \mathbf{B}[0, 1/r]$. This proves that $\mathbf{B}(\infty, r) \cap \mathbf{B}(\hat{P}_j, r) = \emptyset$ for all $j \geq 2$.

Now let $j, j' \geq 2$ and $j' \neq j$. For any $\hat{P} \in \mathbf{B}(\hat{P}_j, r)$, one has $|\hat{P} - \hat{P}_j|_p < r < 1$ and $|\hat{P}_j - \hat{P}_{j'}|_p = 1$ by the previous paragraph and so $|\hat{P} - \hat{P}_{j'}|_p = 1 > r$. This shows $\hat{P} \notin \mathbf{B}(\hat{P}_{j'}, r)$. This proves the disjointness.

The proof for the rest of the lemma follows directly from [19, Theorem 4.7] (see also [12]). $\square$

Every element in the Ω_a-space $\mathcal{H}(\mathbf{A}_r)$ can be uniquely represented as $\sum_{i\geq 0} c_{1,i}X^i + \sum_{j=2}^{\ell}\sum_{i\geq 1} c_{j,i}(X-\hat{P}_j)^{-i}$ where $c_{j,i}\in\Omega_a$ and $\forall j\geq 1, \lim_{i\to\infty}\frac{|c_{j,i}|_p}{r^i}=0$. For simplicity, we write $X_1 := X$ and $X_j := (X-\hat{P}_j)^{-1}$ for $1\leq j\leq \ell$. Then the Ω_a-space $\mathcal{H}(\mathbf{A}_r)$ has a natural monomial basis $\vec{b}_{\rm unw} := \{1, X_1^i, X_2^i, \ldots, X_\ell^i\}_{i\geq 1}$. In Theorem 3.5 we shall use a *weighted* basis $\vec{b}_{\rm w}$ (note that neither $\vec{b}_{\rm unw}$ nor $\vec{b}_{\rm w}$ is an orthonormal basis):

Lemma 2.2. *Let $r\in|\Omega_a|_p$ and $0<r<1$. For each $1\leq j\leq \ell$, let $Z_j := \gamma^{\frac{1}{d_j}}X_j$. Then $\vec{b}_{\rm w} := \{1, Z_1^i, \ldots, Z_\ell^i\}_{i\geq 1}$ forms a basis of $\mathcal{H}(\mathbf{A}_r)$ over Ω'_a.*

Proof. Obvious by Lemma 2.1 and remarks preceding the lemma. □

2.3. **The U_p operator.** For any $s\in p^{\mathbb{Z}_{\geq 0}}$ and for any $\xi(X)\in\mathcal{H}(\mathbf{A}_{r,s})$, let U_p be the map defined by $(U_p\xi)(X) := 1/p\cdot\sum_{Z^p=X}\xi(Z)$ from $\mathcal{H}(\mathbf{A}_{r,s})$ to $\mathcal{H}(\mathbf{A}_{r^p,sp})$. Similarly let $(U_q\xi)(X) := 1/q\cdot\sum_{Z^q=X}\xi(Z)$. This subsection was influenced by the spirit in [8] (in particular Section 3.5). This subsection aims to prove Theorem 2.4 with the following lemma.

Lemma 2.3. *Let $X\in\Omega_a$, $\hat{P}\in\mathcal{O}_a$, and $s\in p^{\mathbb{Z}_{>0}}$. If $r\geq p^{-\frac{p}{p-1}}$ then (1) $|X-\hat{P}|_p > r^{\frac{1}{s}}$ implies $|X^s-\hat{P}^s|_p = |X-\hat{P}|_p^s > r$, and (2) $|X^s-\hat{P}^s|_p > r$ implies $|X^s-\hat{P}^s|_p = |X-\hat{P}|_p^s$.*

Proof. (1) If $\hat{P}=0$ then the lemma is trivial. We assume $\hat{P}\neq 0$ for the rest of the proof. Write $s=p^k$ for $k\geq 1$. We shall use induction on k. We first prove the case $k=1$ for both statements. That is, $|X^p-\hat{P}^p|_p > p^{-\frac{p}{p-1}}$ implies that $|X-\hat{P}|_p > p^{-\frac{1}{p-1}}$ which in turn implies that $|X^p-\hat{P}^p|_p = |X-\hat{P}|_p^p$.

Write $Y := X-\hat{P}$. Then $X^p-\hat{P}^p = (Y+\hat{P})^p-\hat{P}^p = Y^p+pG$ where $G=\sum_{m=1}^{p-1}(\binom{p}{m}/p)Y^{p-m}\hat{P}^m\in Y\mathbb{Z}[\hat{P},Y]$. If $|Y|_p\leq p^{-\frac{1}{p-1}}$, that is, ${\rm ord}_pY\geq\frac{1}{p-1}$, then $|pG|_p\leq|pY|_p\leq p^{-\frac{p}{p-1}}$. Since $|\hat{P}|_p\leq 1$, one has ${\rm ord}_pG\geq{\rm ord}_pY$. Thus $|X^p-\hat{P}^p|_p\leq\max(|Y^p|_p,|pG|_p)\leq p^{-\frac{p}{p-1}}$. Contradiction, so we have $|Y|_p > p^{-\frac{1}{p-1}}$. This implies that ${\rm ord}_pY^{-i} > -\frac{i}{p-1}$ for any $i\in\mathbb{Z}$. By the triangle inequality ${\rm ord}_pG/Y^p\geq\min_{1\leq i\leq p-1}{\rm ord}_pY^{-i} > -1$. Hence ${\rm ord}_p(pG) > {\rm ord}_pY^p$. Again by the triangle inequality we have $|X^p-\hat{P}^p|_p = |Y^p|_p = |X-\hat{P}|_p^p$.

(2) Suppose it holds for $s=p^{k-1}$. By assumption, $|X^{p^k}-\hat{P}^{p^k}|_p = |(X^p)^{p^{k-1}}-(\hat{P}^p)^{p^{k-1}}|_p > p^{-\frac{p}{p-1}}$. By inductive argument one has $|X^{p^k}-\hat{P}^{p^k}|_p = |X^p-\hat{P}^p|_p^{p^{k-1}}$ and so $|X^p-\hat{P}^p|_p > p^{-\frac{p}{p^{k-1}(p-1)}}\geq p^{-\frac{p}{p-1}}$. The latter implies that $|X^p-\hat{P}^p|_p = |X-\hat{P}|_p^p$, again by induction. Therefore, one has $|X^{p^k}-\hat{P}^{p^k}|_p = |X-\hat{P}|_p^{p^k}$, as we desire. Now suppose $|X-\hat{P}|_p > r^{\frac{1}{p^k}}$. Then $|X-\hat{P}|_p > p^{-\frac{1}{p-1}}$ and so $|X^p-\hat{P}^p|_p = |X-\hat{P}|_p^p > r^{\frac{1}{p^{k-1}}}$. Then we use induction argument to get

$$|X^{p^k}-\hat{P}^{p^k}|_p = |(X^p)^{p^{k-1}}-(\hat{P}^p)^{p^{k-1}}|_p = |X^p-\hat{P}^p|_p^{p^{k-1}} = |X-\hat{P}|_p^{p^k}.$$

This finishes our proof. □

Theorem 2.4. *Let $r\in|\Omega_a|_p$ and $p^{-\frac{1}{p^{a-1}(p-1)}} < r < 1$. Let $s\in p^{\mathbb{Z}_{\geq 0}}$. Then $U_p\mathcal{H}(\mathbf{A}_{r,s})\subseteq\mathcal{H}(\mathbf{A}_{r^p,sp})$. Then $U_q=U_p^a$ and $U_q\mathcal{H}(\mathbf{A}_r)\subseteq\mathcal{H}(\mathbf{A}_{r^q})$.*

Proof. 1) We shall demonstrate a proof for the case $s=1$ since the general case is very similar. Let $\xi \in \mathcal{H}(\mathbf{A}_r)$.

Firstly, we show that $U_p\xi$ defines a function on the affinoid $\mathbf{A}_{r^p,p}$. It suffices to show that $Z^p = X \in \mathbf{A}_{r^p,p}$ implies that $Z \in \mathbf{A}_r$. Indeed, for every $2 \le j \le \ell$ one has $|Z^p - \hat{P}_j^p|_p \ge r^p > p^{-\frac{p}{p-1}}$ by hypothesis. By Lemma 2.3 one has $|Z-\hat{P}_j|_p^p = |Z^p - \hat{P}_j^p|_p \ge r^p$. That is, $|Z-\hat{P}_j|_p \ge r$. On the other hand, by $|Z^p|_p \le 1/r^p$, one has $|Z|_p \le 1/r$. This proves our claim.

Secondly we show that $U_p\xi \in \mathcal{H}(\mathbf{A}_{r^p,p})$. Our proof below follows [8, Lemma on page 40]. Before we start, an easy fact is prepared:

$$\sup_{X\in\mathbf{A}_{r^p,p}} |(U_p\xi)(X)|_p \quad \le \quad p\cdot \sup_{X\in\mathbf{A}_r} |\xi(X)|_p. \tag{7}$$

Let Tr denote the trace map from $\Omega_a(Z)$ to $\Omega_a(X)$ where Z is a function with $Z^p = X$. If $\xi \in \Omega_a(X)$, then by definition $U_p\xi = \frac{1}{p}\cdot \mathrm{Tr}\ \xi(X)$. This shows that U_p maps $\Omega_a(X)$ to itself and by (7), if ξ has no pole in $\mathbf{A}_r$ then $U_p\xi$ has no pole in $\mathbf{A}_{r^p,p}$. Thus U_p restricts to a mapping $\Omega_a(X)\cap\mathcal{H}(\mathbf{A}_r) \longrightarrow \Omega_a(X)\cap\mathcal{H}(\mathbf{A}_{r^p,p})$, which is continuous relative to the supremum norms. Since $\xi\in\mathcal{H}(\mathbf{A}_r)$, one gets that ξ may be uniformly approximated on $\mathbf{A}_r$ by elements of $\Omega_a(X)\cap\mathcal{H}(\mathbf{A}_r)$ and so by (7) again $U_p\xi$ can be uniformly approximated on $\mathbf{A}_{r^p,p}$ by elements of $\Omega_a(X)\cap\mathcal{H}(\mathbf{A}_{r^p,p})$. This completes the proof of the assertion about U_p.

2) Let $Z^q = X \in \mathbf{A}_{r^q}$ for $r > p^{-\frac{1}{p^{a-1}(p-1)}}$. For $2\le j\le \ell$, one has $|Z^q - \hat{P}_j^q|_p \ge r^q$; so $|Z-\hat{P}_j|_p = |Z^q-\hat{P}_j^q|_p^{1/q} \ge r$. One also observes that $|Z^q|_p \le 1/r^q$ implies that $|Z|_p \le 1/r$. This proves that $Z\in\mathbf{A}_r$. This proves that $U_q\xi$ is a function on $\mathbf{A}_{r^q,q}$. As $\hat{P}_j^q = \hat{P}_j$ for all j one has $\mathbf{A}_{r^q,q} = \mathbf{A}_{r^q}$, it follows that $U_q\xi$ is defined over $\mathbf{A}_{r^q}$. □

2.4. **Push-forward maps and Dwork's splitting functions.** In the previous subsection we have defined the U_p and U_q operators on suitable p-adic Banach spaces. It remains to define the "Dwork's splitting function" to finish the process of defining the Frobenius map. Let τ be a lift of the Frobenius endomorphism $c\mapsto c^p$ of $\mathbb{F}_{p^a}$ to Ω_a which fixes Ω_1. Thus τ generates $\mathrm{Gal}(\Omega_a/\Omega_1)$. Let $\mathbf{A}_{r,s}^\tau$ denote the image of $\mathbf{A}_{r,s}$ under τ.

Since one may have a pole $\hat{P}_j$ other than 0 and ∞, one encounters the following problem: for any $\xi(X)\in\mathcal{H}(\mathbf{A}_r)$ its image $\xi^\tau(X)$ does not lie in $\mathcal{H}(\mathbf{A}_r^\tau)$ anymore. So the naive generalization of Dwork's splitting function does not work. This prompts us to define some push-forward maps.

Define a map of p-adic Banach spaces

$$\begin{array}{ccc} \mathcal{H}(\mathbf{A}_{r,s}) & \xrightarrow{\tau_*} & \mathcal{H}(\mathbf{A}_{r,s}^\tau) \\ \xi & \mapsto & \tau\circ\xi\circ\tau^{-1}. \end{array}$$

For any $k\in\mathbb{Z}_{\ge 0}$, one has $\tau_*(\mathcal{H}(\mathbf{A}_{r,p^k})) = \mathcal{H}(\mathbf{A}_{r,p^k}^\tau) = \mathcal{H}(\mathbf{A}_{r,p^{k+1}})$. As a simple example, for $B\in\Omega_a$ and a Teichmüller lift $\hat{P}$ in Ω_a with $\xi(X) = \frac{B}{X-\hat{P}} \in \mathcal{H}(\mathbf{A}_r)$ we have $(\tau_*\xi)(X) = \frac{\tau(B)}{X-\tau(\hat{P})} = \frac{\tau(B)}{X-\hat{P}^p}$. On the other hand, one may check routinely that τ_*^k commutes with U_p for any $k\in\mathbb{Z}$.

For any $\hat{f}(x)$ (fixed in Section 2.1), and for every $1 \le j \le \ell$, let

$$F_j(X_j) \quad := \quad \prod_{i=1}^{d_j} E(\gamma a_{j,i} X_j^i) \tag{8}$$

where we recall that $E(X)$ is the Artin-Hasse exponential function and γ is the root of $\log E(X)$ with $\mathrm{ord}_p \gamma = \frac{1}{p-1}$. We now induce our new splitting functions:

$$F(X) := \prod_{j=1}^{\ell} F_j(X_j); \qquad F_{[a]}(X) := \prod_{k=0}^{a-1} (\tau_*^k F)(X^{p^k}). \tag{9}$$

Lemma 2.5. *Let $k \ge 1$ be any integer. Let $r \in |\Omega_a|_p$ and $p^{-\frac{p}{p-1}} < r < 1$. Then for any $\xi(X) \in \mathcal{H}(\mathbf{A}_r^{\tau^k})$ one has $\xi(X^{p^k}) \in \mathcal{H}(\mathbf{A}_{r^{1/p^k}})$.*

Proof. It suffices to show that $|X - \hat{P}_j|_p \ge r^{1/p^k}$ implies that $|X^{p^k} - \hat{P}_j^{p^k}|_p \ge r$. This follows from Lemma 2.3 immediately. $\square$

Theorem 2.6. *Let $d_0 := \max_{1 \le j \le \ell} d_j$. Let $r \in |\Omega_a|_p$ and $p^{-\frac{1}{d_0 p^{a-1}(p-1)}} < r < 1$. Then $F(X) \in \mathcal{H}(\mathbf{A}_{r^{p^{a-1}}}) \subseteq \mathcal{H}(\mathbf{A}_r)$ and $F_{[a]}(X) \in \mathcal{H}(\mathbf{A}_r)$.*

Proof. Write $r_j := p^{-\frac{1}{d_j(p-1)}}$. Write $F_j(X) = \sum_{n=0}^{\infty} F_{j,n} X^n$ over $\mathcal{O}_a$. Note that $F_j(X)$'s convergence radius is $\liminf_n |F_{j,n}|_p^{-1/n} = p^{\liminf_n \mathrm{ord}_p F_{j,n}/n} \ge 1/r_j$. (See Lemma 3.2.) By hypothesis, one has $r^{p^{a-1}} > r_j$ for every j, so $F_1(X), F_j((X - \hat{P}_j)^{-1}) \in \mathcal{H}(\mathbf{A}_{r^{p^{a-1}}})$ for $2 \le j \le \ell$. Hence $F(X) \in \mathcal{H}(\mathbf{A}_{r^{p^{a-1}}})$.

Then $(\tau_*^k F)(X) \in \mathcal{H}(A_{r^{p^{a-1}}}^{\tau^k})$ for every $0 \le k \le a-1$. Our hypothesis implies that $p^{-\frac{p}{p-1}} < r^{p^{a-1}} < 1$. So one may apply Lemma 2.5 and gets, $(\tau_*^k F)(X^{p^k}) \in \mathcal{H}(\mathbf{A}_{r^{p^{a-1-k}}}) \subseteq \mathcal{H}(\mathbf{A}_r)$ for every $0 \le k \le a-1$. Therefore, their product $F_{[a]}(X)$ lies in $\mathcal{H}(\mathbf{A}_r)$ as well. $\square$

2.5. **The trace formula of α_a.**

For the rest of the paper we assume

$$\boxed{r \in |\Omega_a|_p \text{ and } p^{-\frac{1}{d_0 p^{a-1}(p-1)}} < r < 1 \text{ where } d_0 = \max_{1 \le j \le \ell} d_j.} \tag{10}$$

This bound of r is to assure that $F_{[a]}(X)$ lies in $\mathcal{H}(\mathbf{A}_r)$. Let $\alpha_a := U_q \circ F_{[a]}(X)$, by which we mean the composition map of U_q with the multiplication map by $F_{[a]}(X)$. Then α_a is a Ω_a-linear map from $\mathcal{H}(\mathbf{A}_r)$ to $\mathcal{H}(\mathbf{A}_{r^q})$ by Theorem 2.4. Composing with the natural restriction map $\mathcal{H}(\mathbf{A}_{r^q}) \to \mathcal{H}(\mathbf{A}_r)$, one observes that α_a defines an endomorphism of $\mathcal{H}(\mathbf{A}_r)$.

Lemma 2.7 (Dwork-Monsky-Reich)**.** *Let $\overline{f} \in \mathbb{A}(\mathbb{F}_q)$. Let r be as in (10), then the Ω_a-linear endomorphism α_a of $\mathcal{H}(\mathbf{A}_r)$ is completely continuous and one has*

$$L(\overline{f}; T) \quad = \quad \frac{\det(1 - T\alpha_a | \mathcal{H}(\mathbf{A}_r))}{\det(1 - Tq\alpha_a | \mathcal{H}(\mathbf{A}_r))}. \tag{11}$$

Proof. Let $\mathcal{H}^\dagger(\mathbf{A}_1) := \bigcup_{0<r<1} \mathcal{H}(\mathbf{A}_r)$. One notes that $\mathcal{H}^\dagger(\mathbf{A}_1)$ is the Monsky-Washnitzer dagger space of $\mathbf{A}_1$. Our assertions then follow from the trace formula of [15] and [18], as explained in [20, Section 6] and see (6.3.11) in particular. (Our hypothesis $g(0) = 0$ was used there). Basically their trace formula says that α_a is completely continuous on $\mathcal{H}^\dagger(\mathbf{A}_1)$ and $\det(1 - T\alpha_a | \mathcal{H}^\dagger(\mathbf{A}_1)) = \det(1 - T\alpha_a | \mathcal{H}(\mathbf{A}_r))$ for any r within our range in (10). Since it is routine to check this, we omit details. $\square$

Remark 2.8. One can also formulate the above trace formula using Berthelot's rigid cohomology theory. See [1] for detailed annotation of Robba's formulation in [20, Section 6].

2.6. **Descent from α_a to α_1.** The results in this subsection are only used in the proof of Theorem 1.1 in Section 4. Below we shall use a subindex in $\det_{\Omega_1}(\cdot)$ or $\det_{\Omega_a}(\cdot)$ to emphasize our consideration of a map over Ω_1-space or Ω_a-space, respectively. We shall omit the base space $\mathcal{H}(\mathbf{A}_r)$ in $\det(\cdot)$ if the context clearly assures that no confusion is possible. The upshot of our argument is to "descent" the α_a map of the Ω_a-space $\mathcal{H}(\mathbf{A}_r)$ to the α_1 map of the Ω_1-space $\mathcal{H}(\mathbf{A}_r)$. This idea appeared initially in [7, Section 7]. In this paper we use $\mathrm{NP}_p(\cdot)$ and $\mathrm{NP}_q(\cdot)$ to denote p-adic and q-adic Newton polygons, respectively. (These should not be confused with $\mathrm{NP}(\overline{f};\mathbb{F}_q)$.) We use $1/a\cdot\mathrm{NP}_p(\cdot)$ to denote the image of $\mathrm{NP}_p(\cdot)$ shrunk by a factor of $1/a$.

Lemma 2.9. *Let $\alpha_1 := \tau_*^{-1}\circ U_p\circ F(X)$. Then α_1 is a completely continuous τ^{-1}-linear map from $\mathcal{H}(\mathbf{A}_r)$ to $\mathcal{H}(\mathbf{A}_{r^p})$ (over Ω_a) and $\alpha_a=\alpha_1^a$ as Ω_1-linear maps. Then*

$$\det_{\Omega_a}(1-T^a\alpha_a)^a \quad = \quad \prod_{k=0}^{a-1}\det_{\Omega_1}(1-T\zeta_a^k\alpha_1), \tag{12}$$

where ζ_a is a primitive a-th root of unity. Then $\mathrm{NP}_q(\det_{\Omega_a}(1-T\alpha_a)) = 1/a\cdot\mathrm{NP}_p(\det_{\Omega_1}(1-T\alpha_1))$.

Proof. As we already remarked at the beginning of Section 2.4, τ_*^{-1} and U_p commute with each other. For any $k\in\mathbb{Z}$, the Ω_1-linear multiplication map of $(\tau_*^kF)(X)$ on $\mathcal{H}(A_r^{\tau^k})$ can be written as $\tau_*^k\circ F(X)\circ\tau_*^{-k}$. On the other hand, for any function $H_k(X)\in\mathcal{H}(\mathbf{A}_r^{\tau^k})$ one has a general identity stating that

$$U_q\circ\prod_{k=0}^{a-1}H_k(X^{p^k}) \quad = \quad \prod_{k=0}^{a-1}U_p\circ H_{a-1-k}(X), \tag{13}$$

where second product is noncommutative and its factors are ordered from left to right as k increases. We retain this notation of noncommutative products for the rest of the paper.

Now apply (13) to $F_{[a]}(X)$ with $H_k(X):=(\tau_*^kF)(X)$. One has

$$\begin{aligned} U_q\circ F_{[a]}(X) &= \prod_{k=0}^{a-1}(U_p\circ\tau_*^{a-1-k}\circ F(X)\circ\tau_*^{-(a-1-k)}) \\ &= \prod_{k=0}^{a-1}(\tau_*^{a-1-k}\circ U_p\circ F(X)\circ\tau_*^{-(a-1-k)}). \end{aligned}$$

By telescoping, one gets $U_q\circ F_{[a]}(X) = (\tau_*^{-1}\circ U_p\circ F(X))^a$. That is, $\alpha_a=\alpha_1^a$.

The proof for α_1 being completely continuous is verbatim for α_a which is already proved in Lemma 2.7. Now it is elementary to see that

$$\det_{\Omega_1}(1-T^a\alpha_1^a) = \prod_{k=0}^{a-1}\det_{\Omega_1}(1-T\zeta_a^k\alpha_1).$$

One may also show as an exercise that (see [3, (41)] for details)

$$\det_{\Omega_a}(1-T\alpha_a)^a = \det_{\Omega_1}(1-T\alpha_a).$$

Combining these two equalities with $\alpha_a = \alpha_1^a$, one obtains (12). The last assertion about Newton polygons follows from the elementary theory of Newton polygons (see [6, Lemma 1.6] and [7, Lemma 7.1]). $\square$

Proposition 2.10. *The slope < 1 part (of horizontal length $d-\ell+1$) of* $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ *is equal to* $\mathrm{NP}_q(\det_{\Omega_a}(1-T\alpha_a) \bmod T^{d-\ell+1})$ *which is equal to*

$$1/a \cdot \mathrm{NP}_p(\det_{\Omega_1}(1-T\alpha_1) \bmod T^{a(d-\ell+1)+1}).$$

The same holds if one replaces Ω_1 and Ω_a by Ω_1' and Ω_a', respectively.

Proof. By (11), one has

$$L(f \bmod \mathcal{P};T)\cdot \det_{\Omega_a}(1-Tq\alpha_a) \quad = \quad \det_{\Omega_a}(1-T\alpha_a). \tag{14}$$

Note that all slopes are greater than or equal to 1 in $\mathrm{NP}_q(\det_{\Omega_a}(1-Tq\alpha_a))$. By the Weil conjectures for (projective) curves, $L(\overline{f};F_q)$ is a degree d polynomial (see (4)) with all slopes in $[0,1]$. The slope-1 part of $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ is precisely of horizontal length $\ell-1$ (see Remark 1.4). Let λ be the biggest slope of $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ that is strictly less than 1. Then the slope $\leq \lambda$ part of $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ is equal to $\mathrm{NP}_q(\det_{\Omega_a}(1-T\alpha_a) \bmod T^{d-\ell+2})$ by (14) and the p-adic Weierstrass preparation theorem (see section [11, IV.4]).

By Lemma 2.9, $1/a\cdot \mathrm{NP}_p(\det_{\Omega_1}(1-T\alpha_1)) = \mathrm{NP}_q(\det_{\Omega_a}(1-T\alpha_a))$. By the previous paragraph, the latter polygon has a vertex point at $T^{d-\ell+1}$, which separates the slope $\leq\lambda$ and slope-1 segments. Hence the former polygon has a corresponding vertex point at $T^{a(d-\ell+1)}$. The upshot is that

$$1/a\cdot \mathrm{NP}_p(\det_{\Omega_1}(1-T\alpha_1) \bmod T^{a(d-\ell+1)+1}) = \mathrm{NP}_q(\det_{\Omega_a}(1-T\alpha_a) \bmod T^{d-\ell+2}).$$

Compiling these two paragraphs, our assertion follows. The last assertion is obvious. $\square$

3. p-adic estimates of L-functions of exponential sums

This section aims to prove Theorem 3.5 whose proof is however very technical, so the reader is recommended to refer to it only when needed. We retain all notations from previous sections, in particular we recall the two bases $\vec{b}_{\mathrm{unw}}$ and $\vec{b}_{\mathrm{w}}$ of $\mathcal{H}(\mathbf{A}_r)$ from Lemma 2.2. For any $c\in\mathbb{R}$ we denote by $\lceil c\rceil$ the least integer greater than or equal to c.

We start with a lemma inspired by a "Dwork's Lemma" in [8]:

Lemma 3.1. *Let $m\geq 1$ and $J\geq 3$. Let r be as in (10) of Section 2.5. Then for any $X\in\mathbf{A}_r$ one has*

$$U_p(X-\hat{P}_J)^{-m} \quad = \quad \sum_{n=\lceil m/p\rceil}^{m} C^{n,m}\hat{P}_J^{np-m}(X-\hat{P}_J^p)^{-n}, \tag{15}$$

where $C^{n,m}\in\mathbb{Z}_p$ with $\mathrm{ord}_p C^{n,m} \geq \frac{np-m}{p-1}-1$. Then $\mathrm{ord}_p C^{n,m}=0$ if and only if $n=\lceil\frac{m}{p}\rceil$. If this is the case, then one has $C^{n,m}\equiv(-1)^{\epsilon-1} \bmod p$ where $\epsilon = m-(n-1)p$.

Proof. By Theorem 2.4 one has $U_p(X-\hat{P}_J)^{-m}\in\mathcal{H}(\mathbf{A}_{r^p,p})$. The first statement of this lemma follows from an analogous verification as presented in Section 5.3 of [8], so we shall omit its proof here. We shall prove our last assertion below.

By [8, Lemma of Section 5.3, page 74], one has $C^{n,m} = \frac{m}{np}\sum_{\vec{i}}\prod_{k=1}^{n}\binom{p}{i_k}$ where $\vec{i} := (i_1, \ldots, i_n)$ ranges in $\mathbb{Z}^n$ with $1 \le i_1, \ldots, i_n \le p$ and $\sum_{k=1}^{n} i_k = m$ (we denote this set of $\vec{i}$ by $\mathcal{I}$). Write $m = (n-1)p + r$ for some $1 \le r \le p$. If $r = p$ then one can easily see that $C^{n,m} = 1$ and our assertion clearly holds. Below we let $1 \le r \le p-1$. We assume additionally that $p > 2$ since if $p = 2$ then one has $m = 2n-1$ and it is easy to check that $\mathrm{ord}_p C^{n,m} = 0$ directly. Write $\varsigma(\vec{i}) := \prod_{k=1}^{n}\binom{p}{i_k}$. For any $1 \le t \le n$ let $\mathcal{I}_t$ be the subset of $\mathcal{I}$ consisting of all $\vec{i}$ with $\mathrm{ord}_p(\varsigma(\vec{i})) = t$. It is clear that $\mathcal{I} = \bigcup_{t=1}^{n}\mathcal{I}_t$ is a partition of $\mathcal{I}$. Since $\mathrm{ord}_p\binom{p}{i_k} = 0$ (resp., $= 1$) if and only if $i_k = p$ (resp., $\neq p$), one gets for any $1 \le t \le n$ that $\vec{i} \in \mathcal{I}_t$ if and only if the $\vec{i}$ contain precisely t non-p components. For each $\vec{i} \in \mathcal{I}_t$ there are actually $\binom{n}{t}$ of them by a permutation of the non-p components among the n components and they have the same $\varsigma(\vec{i})$. Let $\mathcal{J}_t$ be the set of all t-tuples $\vec{j} := (j_1, \ldots, j_t)$ with $1 \le j_1, \ldots, j_t \le p-1$ and $\sum_{k=1}^{t} j_k = (t-1)p + r$. Then one gets

$$
\begin{aligned}
C^{n,m} &= \sum_{t=1}^{n}\sum_{\vec{i}\in\mathcal{I}_t}\frac{m}{np}\varsigma(\vec{i}) = \sum_{t=1}^{n}\sum_{\vec{j}\in\mathcal{J}_t}\frac{m}{np}\binom{n}{t}\varsigma(\vec{j}) \\
&= \frac{m\binom{p}{r}}{p} + \sum_{t=2}^{n}\sum_{\vec{j}\in\mathcal{J}_t}\frac{m}{np}\binom{n}{t}\varsigma(\vec{j}).
\end{aligned} \tag{16}
$$

One easily observes that the first summand is a p-adic unit. Now we claim that for any $t \ge 2$ and $\vec{j} \in \mathcal{J}_t$ one has $\mathrm{ord}_p(\frac{m}{np}\binom{n}{t}\varsigma(\vec{j})) \ge 1$. Indeed, one has for some $u \in Z_p$ that

$$\frac{m}{np}\binom{n}{t}\varsigma(\vec{j}) = \frac{m}{np}\left(\frac{n}{t}\binom{n-1}{t-1}\right)(p^t u) = um\binom{n-1}{t-1}\frac{p^{t-1}}{t}.$$

It is easy to observe that for any $t \ge 2$ and $p > 2$ one has $\mathrm{ord}_p t \le t-2$. This proves our claim above. By (16), we get that $C^{n,m}$ is a p-adic unit. □

The computation of $\alpha_1 = \tau_*^{-1} \circ U_p \circ F(X)$ uses the observation that τ_*^{-1} and U_p respect the Mittag-Leffler decomposition while the multiplication map $F(X)$ does not. For $1 \le j \le \ell$ and for any $\xi(X) \in \mathcal{H}(\mathbf{A}_r)$, let $\xi(X)_{\hat{P}_j}$ denote the j-th component in the p-adic Mittag-Leffler decomposition as in Lemma 2.1. We recall our notation $X_1 = X$ and $X_j = (X - \hat{P}_j)^{-1}$ for $2 \le j \le \ell$.

Now we recall certain properties of $F_j(X_j) = \sum_{n=0}^{\infty} F_{j,n}X_j^n \in \mathcal{O}_a[[X_j]]$ (see proof in, for instance, [29, Section 1]).

Lemma 3.2. *For any $1 \le j \le \ell$ and $n \ge 0$ one has $\mathrm{ord}_p F_{j,n} \ge \frac{\lceil \frac{n}{d_j} \rceil}{p-1}$ where the equality holds if $d_j | n$ and $\frac{n}{d_j} \le p-1$. In particular, $\mathrm{ord}_p F_{j,n} > 0$ for any $n > 0$.*

Lemma 3.3 (Key computational lemma)**.** *(1) If $\xi(X) \in \mathcal{H}(\mathbf{A}_r)$ is given by its Laurent expansion at $\hat{P}_J$, that is $\xi(X) = \sum_{n=-\infty}^{\infty} B_n X_J^n$ for some $B_n \in \Omega_a$, then $(\xi(X))_{\hat{P}_J} = \sum_{n=0}^{\infty} B_n X_J^n$, and $B_0 = 0$ if $\hat{P}_J \neq \infty$.*

(2) Recall $F(X)$ and $C^{n,m}$ from (9) and (15) respectively. For all $i \ge 0$ write $(F(X)X_J^i)_{\hat{P}_{J_1}} = \sum_{n=0}^{\infty} H_{J_1,J}^{n,i} X_{J_1}^n$ for some $H_{J_1,J}^{n,i} \in \Omega_a$. Then

$$(\alpha_1 X_J^i)_{\hat{P}_{J_1}} = \sum_{n=0}^{\infty} B_{J_1,J}^{n,i} X_{J_1}^n \in \Omega_a[[X_{J_1}]]$$

where

$$B_{J_1,J}^{n,i} := \begin{cases} \tau^{-1} H_{J_1,J}^{np,i} & \textit{for } J_1 = 1,2 \\ \sum_{m=n}^{np} C^{n,m} \hat{P}_{J_1}^{n-mp^{a-1}} \tau^{-1} H_{J_1,J}^{m,i} & \textit{for } J_1 \geq 3. \end{cases}$$

Proof. Part (1) is a simple corollary of the remarks preceding the lemma and Lemmas 2.1 and 3.1. The rest are routine consequences. □

For any integers $s \geq 0$ and $t \geq 1$ we use $\mathcal{C}(s,t)$ to denote the condition that $t|s$ and $0 \leq \frac{s}{t} \leq p-1$ are satisfied (e.g., the condition in Lemma 3.2 is $\mathcal{C}(n, d_j)$). We claim the following:

$$|H_{J_1,J}^{n,i}|_p \leq \begin{cases} p^{-\frac{n-i}{d_{J_1}(p-1)}} & \text{if } J_1 = J \\ p^{-\frac{n+i}{d_{J_1}(p-1)}} & \text{if } J_1 = 1 \neq J \\ p^{-\frac{n}{d_{J_1}(p-1)}} & \text{if } J_1 = 2 \neq J. \end{cases} \tag{17}$$

Furthermore, the equalities hold if and only if additional conditions $\mathcal{C}(n-i, d_{J_1})$, $\mathcal{C}(n+i, d_{J_1})$, $\mathcal{C}(n, d_{J_1})$ hold, respectively.

A proof for the case $J = J_1$ is sketched below and proofs for other cases are omitted as they are formal and similar. Let $\vec{n} := (n_1, \ldots, n_\ell) \in \mathbb{Z}_{\geq 0}^\ell$. Then one notices that for $J = 1$

$$H_{1,1}^{n,i} = \sum \left(F_{1,n_1} \prod_{j \neq 1} \left(\sum_{m_j=0}^{n_j} F_{j,m_j} \binom{n_j - 1}{m_j - 1} \hat{P}_j^{n_j - m_j} \right) \right),$$

and for $J \geq 2$,

$$\begin{aligned} H_{J,J}^{n,i} = & \sum \left(F_{J,n_J} \left(\sum_{m_1=n_1}^{\infty} F_{1,m_1} \binom{m_1}{n_1} \hat{P}_J^{m_1-n_1} \right) \right. \\ & \left. \cdot \prod_{j \neq 1,J} \left(\sum_{m_j=0}^{\infty} F_{j,m_j} (-1)^{m_j} \binom{n_j + m_j - 1}{m_j - 1} (\hat{P}_j - \hat{P}_J)^{-(n_j+m_j)} \right) \right) \end{aligned}$$

where the sums both range over all $\vec{n} \in \mathbb{Z}_{\geq 0}^\ell$ such that $n - i = n_J - \sum_{j \neq J} n_j$.

From the above, one observes that $\mathrm{ord}_p H_{J_1,J}^{n,i}$ is greater than or equal to the minimal valuation among the $\vec{n}$-summand in its formula as $\vec{n}$ varies in its domain. Each $\vec{n}$-summand is the product of ℓ elements in $\mathcal{O}_a$, so its valuation is equal to the sum of the valuations of these ℓ elements in $\mathcal{O}_a$. It is easy to observe that $\mathrm{ord}_p H_{J_1,J}^{n,i} \geq \min_{n_1} (\mathrm{ord}_p F_{1,n_1}) \geq \frac{n-i}{d_{J_1}(p-1)}$ as the minimum is taken over all $n_1 = n - i + \sum_{j=2}^\ell n_j$. Moreover, if $\mathcal{C}(n-i, d_{J_1})$ holds then by Lemma 3.2 the minimal is uniquely achieved at $\vec{n} = (n-i, 0, \ldots, 0)$ and the equality holds. Conversely, suppose the equality in (17) holds. It can be easily seen that $H_{J_1,J}^{n,i}$ lies in $\mathcal{O}_a$, in which p has ramification index $p-1$ over $\mathbb{Z}$, so $\mathrm{ord}_p H_{J_1,J}^{n,i} = \frac{n-i}{d_{J_1}(p-1)}$ lies in $\frac{\mathbb{Z}}{p-1}$. Thus $\mathcal{C}(n-i, d_{J_1})$ holds. This proves our claim in (17).

Theorem 3.4 (Unweighted estimates)**.** *Let $B_{J_1,J}^{n,i} \in \mathcal{O}_a$ be as in Lemma 3.3(2).*

(1) For $J_1 = 1, 2$ and for $n, i \geq 0$ one has

$$|B^{n,i}_{J_1,J}|_p \leq \begin{cases} p^{-\frac{np-i}{d_{J_1}(p-1)}} & \text{if } J_1 = J \\ p^{-\frac{np+i}{d_{J_1}(p-1)}} & \text{if } J_1 = 1 \neq J \\ p^{-\frac{np}{d_{J_1}(p-1)}} & \text{if } J_1 = 2 \neq J. \end{cases}$$

The equalities hold if and only if the additional conditions $\mathcal{C}(np - i, d_{J_1}), \mathcal{C}(np + i, d_{J_1})$, and $\mathcal{C}(np, d_{J_1})$ hold, respectively.

(2) For $J_1 \geq 3$ and for $n \geq 1, i \geq 0$ one has

$$|B^{n,i}_{J_1,J}|_p \leq \begin{cases} p^{-\frac{(n-1)p-(i-1)}{d_{J_1}(p-1)}} & \text{if } J_1 = J \\ p^{-\frac{(n-1)p+1}{d_{J_1}(p-1)}} & \text{if } J_1 \neq J. \end{cases}$$

For $d_{J_1} \geq 2$ the equalities hold if additional conditions $\mathcal{C}((n-1)p-(i-1), d_{J_1})$ and $\mathcal{C}((n-1)p + 1, d_{J_1})$ hold, respectively.

Proof. If $J_1 = 1, 2$ one has $|B^{n,i}_{J_1,J}|_p = |H^{np,i}_{J_1,J}|_p$ by Lemma 3.3. Combining this with (17), part (1) follows immediately. We are left to prove part (2). Assume $J_1 \geq 3$ from now on. We shall outline a proof for the case $J = J_1$: Let n, i be fixed in their appropriate ranges. By Lemma 3.3(2), one has that $|B^{n,i}_{J_1,J}|_p \leq \max_{n\leq m\leq np}(|H^{m,i}_{J_1,J}C^{n,m}|_p)$ and the equality holds if the maximum is unique. Pick $m_0 := (n - 1)p + 1$, then one has two cases: (a) For any $m_0 < m \leq np$ one has $|H^{m,i}_{J_1,J}|_p < p^{-\frac{m_0-i}{d_{J_1}(p-1)}}$ by (17); (b) Let $n \leq m \leq m_0$. The function $c(m) := \frac{m-i}{d_{J_1}(p-1)} + (\frac{np-m}{p-1} - 1) = \frac{np}{p-1} - \frac{i}{d_{J_1}(p-1)} - 1 - m\frac{d_{J_1}-1}{d_{J_1}(p-1)}$ has its minimum $c(m_0)$. If $d_{J_1} \geq 2$ then this minimum is unique. By (17) and Lemma 3.1 one has

$$\max_{n\leq m\leq m_0} |H^{m,i}_{J_1,J}C^{n,m}|_p \leq p^{-c(m_0)} = p^{-\frac{m_0-i}{d_{J_1}(p-1)}} = p^{-\frac{(n-1)p-(i-1)}{d_{J_1}(p-1)}}.$$

Combining (a) and (b) one gets the desired upper bound for $|B^{n,i}_{J_1,J}|_p$. Suppose $d_{J_1} \geq 2$ and $\mathcal{C}(m_0 - i, d_{J_1})$ holds. Then the maximum is achieved uniquely at m_0 by Lemma 3.1. Combining the above, we have proved part (2) for the case $J_1 = J$. Since other cases are similar and we omit them here, and finally we conclude the proof to our theorem. □

Theorem 3.5 (Weighted estimates). *Write $(\alpha_1 Z^i_J)_{\hat{P}_{J_1}} = \sum_{n=0}^{\infty} C^{n,i}_{J_1,J} Z^n_{J_1}$ in $\Omega'_a[[Z_{J_1}]]$. Then $C^{n,i}_{J_1,J} = B^{n,i}_{J_1,J}\gamma^{\frac{i}{d_J}-\frac{n}{d_{J_1}}}$ and $\mathrm{ord}_p C^{n,i}_{J_1,J} = \mathrm{ord}_p B^{n,i}_{J_1,J} + \frac{1}{p-1}(\frac{i}{d_J} - \frac{n}{d_{J_1}})$ where $B^{n,i}_{J_1,J}$ is as in Lemma 3.3 (2).*

(1) For $J_1 = 1, 2$ and $n, i \geq 0$ one has

$$\mathrm{ord}_p C^{n,i}_{J_1,J} \geq \begin{cases} \frac{n}{d_{J_1}} & \text{if } J_1 = J \\ \frac{n}{d_{J_1}} + \frac{i}{p-1}(\frac{1}{d_{J_1}} + \frac{1}{d_J}) & \text{if } J_1 = 1 \neq J \\ \frac{n}{d_{J_1}} + \frac{i}{(p-1)d_J} & \text{if } J_1 = 2 \neq J. \end{cases} \tag{18}$$

The equalities hold if and only if $\mathcal{C}(np - i, d_{J_1}), \mathcal{C}(np + i, d_{J_1})$ and $\mathcal{C}(np, d_{J_1})$ hold, respectively.

(2) For $J_1 \geq 3$ and $n \geq 1, i \geq 0$ one has

$$\mathrm{ord}_p C^{n,i}_{J_1,J} \geq \begin{cases} \frac{n-1}{d_{J_1}} & \text{if } J_1 = J \\ \frac{n-1}{d_{J_1}} + \frac{i}{(p-1)d_J} & \text{if } J_1 \neq J. \end{cases} \tag{19}$$

For $d_{J_1}\geq 2$ the equalities hold if conditions $\mathcal{C}((n-1)p-(i-1),d_{J_1}),\mathcal{C}((n-1)p+1,d_{J_1})$ hold, respectively.

Proof. The first statement is clear by a simple calculation. Parts (1) and (2) follow from Lemma 3.2 and parts (1) and (2) of Theorem 3.4, respectively. □

Notations: let $j,j_1\geq 3$, $j\neq j_1$, $n,i\geq 1$. We put row minimal p-adic valuation in boxes.

$\mathrm{ord}_p(\cdot)\geq$	$Z^i_{J=1}$	$Z^i_{J=2}$	$Z^i_{J=j}$	$Z^i_{J=j_1}$
$Z^n_{J_1=1}$	$\boxed{\frac{n}{d_1}}$	$\frac{n}{d_1}+\frac{i/d_1+i/d_2}{p-1}$	$\frac{n}{d_1}+\frac{i/d_1+i/d_j}{p-1}$	$\frac{n}{d_1}+\frac{i/d_1+i/d_{j_1}}{p-1}$
$Z^n_{J_1=2}$	$\frac{n}{d_2}+\frac{i/d_1}{p-1}$	$\boxed{\frac{n}{d_2}}$	$\frac{n}{d_2}+\frac{i/d_j}{p-1}$	$\frac{n}{d_2}+\frac{i/d_{j_1}}{p-1}$
$Z^n_{J_1=j}$	$\frac{n-1}{d_j}+\frac{i/d_1}{p-1}$	$\frac{n-1}{d_j}+\frac{i/d_2}{p-1}$	$\boxed{\frac{n-1}{d_j}}$	$\frac{n-1}{d_j}+\frac{i/d_{j_1}}{p-1}$
$Z^n_{J_1=j_1}$	$\frac{n-1}{d_{j_1}}+\frac{i/d_1}{p-1}$	$\frac{n-1}{d_{j_1}}+\frac{i/d_2}{p-1}$	$\frac{n-1}{d_{j_1}}+\frac{i/d_j}{p-1}$	$\boxed{\frac{n-1}{d_{j_1}}}$

TABLE 1. Lower bounds for $\mathrm{ord}_p C^{n,i}_{J_1,J}$ in matrix of α_1

4. NEWTON POLYGON LIES OVER HODGE POLYGON

Our proof of Theorem 1.1 consists of three parts. The first two parts are in the spirit of Dwork (see [7, Section 7] or [3, Lemma 2]) after a simple reduction. The third part uses Wan's [28, Theorem 2.4].

We queue up the numbers in (1) in nondecreasing order. For any $i\geq 1$, let m_i be the i-th in this queue. For any $k\geq 1$, let $c_k:=\sum_{i=1}^k m_i$ and set $c_0=0$. It is by elementary arithmetic of Newton polygons that $\mathrm{HP}(\mathbb{A})$ is equal to the connecting graph of $\{(k,c_k)\}_{0\leq k\leq d}$ on $\mathbb{R}^2$.

$\boxed{\text{Part 1. Newton polygon of } \alpha_1 \text{ over } \Omega'_a.}$ From now on let $\mathbf{M}$ be the (infinite) matrix representing the α_1 action on Ω'_a-space $\mathcal{H}(\mathbf{A}_r)$ with respect to the basis $\vec{b}_{\mathrm{w}}$. (See Table 1.) Write

$$\det(1-T\mathbf{M})=1+\sum_{k=1}^{\infty}C_kT^k\in\mathcal{O}'_a[[T]]. \tag{20}$$

Take the minimal p-adic valuation of all entries in each row, and put them in a nondecreasing order. For any $i\geq 1$ let $m_i(\mathbf{M})$ denote the i-th smallest row p-adic valuation of $\mathbf{M}$ (counting multiplicity). For every $k\geq 1$ let $c_k(\mathbf{M}):=\sum_{i=1}^k m_i(\mathbf{M})$. By Theorem 3.5, one has

$$\mathrm{ord}_pC^{n,i}_{J_1,J}\geq\begin{cases}\frac{n}{d_{J_1}} & \text{for } J_1=1,2 \text{ and } n,i\geq 0\\ \frac{n-1}{d_{J_1}} & \text{for } J_1\geq 3 \text{ and } n\geq 1,i\geq 0.\end{cases} \tag{21}$$

This implies that $m_i(\mathbf{M})\geq m_i$ for all $1\leq i\leq d-\ell+1$. Thus by arithmetic of Newton polygons (see [6, Lemma 1.6]) and Fredholm theory (see [23, Proposition 7 and its proof]) one has that $\mathrm{NP}_q(\det(1-T\mathbf{M}) \bmod T^{d-\ell+2})$ lies above the connecting graph of $\{(k,c_k)\in\mathbb{R}^2\}_{0\leq k\leq d-\ell+1}$. The latter is precisely $\mathrm{HP}(\mathbb{A})$ as remarked earlier.

$\boxed{\text{Part 2. Newton polygon of } \alpha_1 \text{ over } \Omega_1'.}$ By the normal basis theorem, there exists $\xi \in \Omega_a$ such that $\vec{\xi} := \{\xi^{\tau^t}\}_{0\le t\le a-1}$ is a basis for Ω_a' over Ω_1'. Let $\mathbf{N}$ be the (infinite) matrix representing α_1 with respect to the basis $\vec{b}_{\mathrm{w},\Omega_1'}$ for $\mathcal{H}(\mathbf{A}_r)$ as Ω_1'-space, where $\vec{b}_{\mathrm{w},\Omega_1'}$ consists of $Z_j^i\xi^{\tau^t}$ for $1\le j\le \ell$, $0\le t\le a-1$ and $i\ge 0$ (where $i=0$ only if $j=1$). Write $\det(1-T\mathbf{N}) = 1+\sum_{k=1}^\infty D_kT^k \in \mathcal{O}_1'[[T]]$. We have $\alpha_1(Z_J^i\xi^{\tau^t}) = \alpha_1(Z_J^i)\xi^{\tau^{t-1}} = \sum_{J_1=1}^{\ell}\sum_{n=0}^\infty C_{J_1,J}^{n,i}Z_{J_1}^n\xi^{\tau^{t-1}}$. Recall the lower bound of $\mathrm{ord}_p(C_{J_1,J}^{n,i})$ given in (21). In the two sequences

$$\{m_1(\mathbf{N}), m_2(\mathbf{N}), \cdots, \cdots, m_{a(d-\ell+1)}(\mathbf{N})\} \tag{22}$$

$$\{\underbrace{m_1,\cdots,m_1}_{a},\cdots,\underbrace{m_{d-\ell+1},\cdots,m_{d-\ell+1}}_{a}\}, \tag{23}$$

one notes that (22) *dominates* (23) in the sense that the i-th term of the former sequence is greater than or equal to that of the latter. Thus $1/a\cdot \mathrm{NP}_p(\det_{\Omega_1}(1-T\alpha_1) \bmod T^{a(d-\ell+1)+1})$ lies above the connecting graph of $\{(k,c_k)\}_{0\le k\le d-\ell+1}$, that is $\mathrm{HP}(\mathbb{A})$. By applying Proposition 2.10, one now concludes that $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ lies over $\mathrm{HP}(\mathbb{A})$.

$\boxed{\text{Part 3. Newton and Hodge coincide if and only if } p\equiv 1 \bmod (\mathrm{lcm}\, d_j).}$ One notes that after permuting our basis $\vec{b}_{\mathrm{w}}$ for $\mathcal{H}(\mathbf{A}_r)$ we can arrive at a matrix $\mathbf{M}$ of α_1 in block form satisfying the hypothesis of [28, Theorem 2.4]. Let $\mathbf{M}_a$ be the matrix representing α_a over Ω_a', then one knows that $\mathbf{M}_a = \mathbf{M}\mathbf{M}^{\tau^{-1}}\cdots\mathbf{M}^{\tau^{-(a-1)}}$. Note that $\mathrm{NP}_q(\det(1-T\mathbf{M}_a))$ and $\mathrm{HP}(\mathbb{A})$ meet at the point with x-coordinate $d-\ell+1$ (see Proposition 2.10). Let $\mathrm{HP}(\mathbb{A})_{<1}$ denote the slope <1 part of $\mathrm{HP}(\mathbb{A})$, which has horizontal length $d-\ell+1$. By [28, Theorem 2.4, Corollary 2.5], one can show that $\mathrm{NP}_q(\det(1-T\mathbf{M}_a) \bmod T^{d-\ell+2}) = \mathrm{HP}(\mathbb{A})_{<1}$ if and only if $\mathrm{NP}_p(\det(1-T\mathbf{M}) \bmod T^{d-\ell+2}) = \mathrm{HP}(\mathbb{A})_{<1}$. That is, it is enough to show $\mathrm{NP}_p(\det(1-T\mathbf{M}) \bmod T^{d-\ell+2}) = \mathrm{HP}(\mathbb{A})_{<1}$.

Let $\mathbf{M}_{<1}$ be the principle submatrix of $\mathbf{M}$ consisting of all $C_{J_1,J}^{n,i}$ with

$$\begin{cases} 0\le n\le d_1-1 & \text{for } J_1=1;\\ 1\le n\le d_2-1 & \text{for } J_1=2;\\ 1\le n\le d_{J_1} & \text{for } J_1\ge 3.\end{cases}$$

One notices that $\mathbf{M}_{<1}$ has $d-\ell+1$ rows in total. By (21), every row of $\mathbf{M}$ outside these $d-\ell+1$ rows has its minimal p-adic valuation greater than or equal to 1. From matrix arithmetic of Fredholm theory, it is not hard to conclude that all segments of $\mathrm{NP}_p(\det(1-T\mathbf{M}))_{<1}$ have to "come from" $\det(1-T\mathbf{M}_{<1})$ in the following sense. In (20) let t be the biggest integer such that $\mathrm{NP}_q(\sum_{k=0}^t C_kT^k)$ has all slopes less than 1, then for all $k\le n$ one has $C_k = \sum_N \pm\det N$ where N ranges over all $k\times k$ principal submatrices of $\mathbf{M}_{<1}$.

Now we assume that $p\equiv 1 \bmod (\mathrm{lcm}\, d_j)$ where j ranges from 1 to ℓ. By Remark 1.4, the slope-0 segment of the Hodge polygon is always achieved. This saves us from considering the corresponding rows in $\mathbf{M}_{<1}$. By Theorem 3.5, for $J_1=1,2$ (resp. $J_1\ge 3$) one has that $C_{J_1,J}^{n,i}$ in the submatrix $\mathbf{M}_{<1}$ achieves its minimal row p-adic value $\frac{n}{d_{J_1}}$ (resp., $\frac{n-1}{d_{J_1}}$) uniquely at $n=i$. In summary, all minimal row p-adic valuations are achieved uniquely on the diagonal of $\mathbf{M}_{<1}$. These minimal row p-adic valuations are precisely the rational numbers less than 1 listed in (1). By arithmetic of Fredholm theorem and analysis in Part 1, $\mathrm{NP}_p(\det(1-T\mathbf{M}) \bmod T^{d-\ell+2}) =$

$\mathrm{NP}_p(\det(1-T\mathbf{M}_{<1}))=\mathrm{HP}(\mathbb{A})_{<1}$. Combining with the above paragraph, we have shown that $\mathrm{NP}_q(\det(1-T\mathbf{M}_a) \bmod T^{d-\ell+2})=\mathrm{HP}(\mathbb{A})_{<1}$. By Proposition 2.10, one concludes that $\mathrm{NP}(\overline{f};\mathbb{F}_q)=\mathrm{HP}(\mathbb{A})$.

Conversely, suppose $\mathrm{NP}(\overline{f};\mathbb{F}_q)$ coincides with Hodge. By the above argument, one has $m_i(\mathbf{M})=m_i$ for all i. Because every minimal row p-adic valuation is achieved only at $J=J_1$ (it lies in the diagonal blocks in the matrix $\mathbf{M}$), the Newton polygon of $\mathbf{M}$ lies above the end-to-end join of those of $\mathbf{M}_j$ for $1\le j\le \ell$ where $\mathbf{M}_j:=\{C_{j,j}^{n,i}\}_{1\le n,i\le d_j-1}$. Thus the Newton polygon of $\mathbf{M}_j$ has to coincide with its Hodge (its Hodge polygon is defined in the obvious sense). By the remark in the second last paragraph in Section 1, one can shift the pole to ∞ so that we may assume $p\not\equiv 1 \bmod d_1$. Since $\mathrm{ord}_p(C_{1,1}^{n,i})=n/d_1$ for some $1\le i_n\le d_1-1$ for all $1\le n\le d_1-1$, by Theorem 3.4, the condition $\mathcal{C}(np-i_n,d_1)$ holds, and it is $np\equiv i_n \bmod d_1$. Since $p\not\equiv 1 \bmod d_1$, one has $n\neq i_n$ for every n. From simple linear algebra, one sees that the first slope of the Newton polygon of $\mathbf{M}_j$ is greater than or equal to $\mathrm{ord}_p(C_{1,1}^{1,i_1})>1/d_1$. A contradiction.

This finishes the proof of Theorem 1.1.

Acknowledgments. I am grateful to Bjorn Poonen and Steve Sperber for communicating to me of their conjecture and for stimulating discussions. I thank Hanfeng Li for many exchanges of ideas and in particular for constructive comments on Section 2.4.

References

[1] Pierre Berthelot: Cohomologie rigide et théorie de Dwork: le cas des sommes exponentielles[Rigid cohomology and Dwork theory: the case of exponential sums], in *Cohomologie p-adique* Société Mathématique de France, Astérisque **119–120**(1984), 17-49 (French).

[2] Pierre Berthelot: Géométrie rigide et cohomologie des variétés algébriques de caractéristique p [Rigid geometry and cohomology of algebraic varieties in characteristic p]. Introductions aux cohomologies p-adiques (Luminy, 1984). Mém. Soc. Math. Fr.(N.S.) **23** (1986), 7–32 (French).

[3] Enrico Bombieri: On exponential sums in finite fields, *Amer. J. Math.* **88** (1966), 71–105.

[4] S. Bosch, U. Guntzer, R. Remment: Non-Archimedean analysis, *Grundlehren der mathematischen Wissenschaften* Vol. **261**, Springer-Verlag, Berlin, 1984.

[5] Richard Crew: Etale p-covers in characteristic p, Compositio Math., **52** (1984), no.1, 31–45.

[6] Bernard Dwork: On the zeta function of a hypersurface. *Inst. Hautes Études Sci. Publ. Math.***12** (1962), 5–68.

[7] Bernard Dwork: On the zeta function of a hypersurface. II. *Ann. of Math.(2)* **80** (1964), 227–299.

[8] Bernard Dwork: *Lectures on p-adic differential equations.* Grundlehren der mathematischen Wissenschaften **253**, Springer-Verlag, New York, 1982.

[9] N. Katz: Slope filtration of F-crystals, *Astérisque* **63** (1979), 113–163.

[10] Neal Koblitz: p-adic variation of the zeta function over families of varieties defined over finite fields, *Comp. Math.* **31** (1975), no.2, 119–218.

[11] Neal Koblitz: p-adic numbers, p-adic analysis, and zeta-functions, (Second edition), *Graduate Texts in Mathematics* **58**. Springer-Verlag, New York, 1984.

[12] Marc Krasner: Prolongement analytique uniforme et multiforme dans les corps valués complets, *Les Tendances Géom. en Algèbre et Théorie des Nombres*, Éditions du Centre National de la Recherche Scientifique, CNRS, Paris, 1966, pp. 97–141 (French).

[13] Hanfeng Li and Hui June Zhu: Asymptotic variation of sheaves of one-variable exponential sums. 2003, preprint, `http://arXiv.org/abs/math.NT/0312423`.

[14] Barry Mazur: Frobenius and the Hodge filtration, Bull. A.M.S. **78** (1972), 653–667.

[15] P. Monsky; Washnitzer: Formal cohomology I, *Ann. of Math.* **88** (1968), 181–217.

[16] P. Monsky: Formal cohomology II, *Ann. of Math.* **88** (1968), 218–238.

[17] P. Monsky: Formal cohomology III, *Ann. of Math.* **93** (1971), 315–343.
[18] Daniel Reich: A p-adic fixed point formula. *Amer. J. Math.* **91** (1969), 835–850.
[19] Philippe Robba: Fonctions analytiques sur les corps valués untramétriques complets, in *Prolongement analytique et algèbres banach ultram etriques*, Astérisque **10**, Société Mathématique de France, Paris, 1973, pp. 109–220 (French).
[20] Philippe Robba: Index of p-adic differential operators III. Application to twisted exponential sums, in *Cohomologie p-adique* Astérisque **119–120** (1984), 191–266.
[21] Philippe Robba: Une Introduction näive aux cohomologies de Dwork, *Soc. Math. Fr. Mémoire (N.S.)* **23** (1986), 61–105 (French).
[22] W.H. Schikhof: Ultrametric calculus. an introduction to p-adic analysis. *Cambridge Studies in Advanced mathematics*, vol.**4**. Cambridge University Press, Cambridge, 1984.
[23] J-P Serre: Endomorphismes complètements continus des espaces de Banach p-adique, *Inst. Hautes. Études Sci. Publ. Math.* **12** (1962), 69–85 (French).
[24] Steven Sperber: Congruence properties of the hyper-Kloosterman sum. *Compositio Math.* **40** (1980), no.1, 3–33.
[25] Steven Sperber: On the p-adic theory of exponential sums, *Amer. J. Math.,* **109** (1986), no.2, 255–296.
[26] Gerard van der Geer; Marcel van der Vlugt: Reed-Muller codes and supersingular curves. I. *Compositio Math.* **84** (1992), no. 3, 333–367.
[27] Gerard van der Geer; Marcel van der Vlugt: On the existence of supersingular curves of given genus. *J. Reine Angew. Math.* **458** (1995), 53–61.
[28] Daqing Wan: Newton polygons of zeta functions and L functions. *Ann. of Math.*, **137** (1993), 249–293.
[29] Hui June Zhu: p-adic variation of L functions of one variable exponential sums, I. *Amer. J. Math.* **125** (2003).
[30] Hui June Zhu: Asymptotic variation of L functions of one-variable exponential sums. *J. Reine Angew. Math.* **572** (2004), 219–233.

Department of mathematics and statistics, McMaster University, Hamilton, ON L8S 4K1, CANADA
E-mail address: `zhu@cal.berkeley.edu`