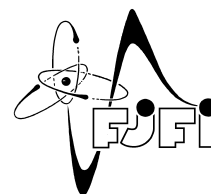




CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering



UNIVERSITY STUDENTS' COMPETITION (SVOČ)

OPTIMAL DISCRETE p -HARDY INEQUALITIES OF HIGHER ORDER

BC. JAKUB WACLAWEK

ABSTRACT. The classical discrete p -Hardy inequality provides a sharp relation between the ℓ^p -norms of a sequence and its discrete derivative on the half-line. In this work, we extend this inequality to higher-order discrete derivatives and establish optimal discrete p -Hardy inequalities of arbitrary order. Specifically, we develop a general method for constructing optimal p -Hardy weights via a parameter sequence subject to certain monotonicity and growth conditions. By applying this abstract framework to a particular choice of the parameter sequence, we obtain explicit formulas for optimal discrete p -Hardy weights of any order. Moreover, by deriving lower bounds for these optimal weights, we find the classical discrete p -Hardy inequalities of any order. As a limiting case of the discrete results, we also obtain their continuous analogs.

This study extends the results of our earlier work, which was submitted to the SVOČ competition in 2024, and, in an expanded form, will constitute the basis of the author's master's thesis.

SUPERVISOR: DOC. ING. FRANTIŠEK ŠTAMPACH, PH.D.

YEAR: 2026

INTRODUCTION

In 1915, the renowned English mathematician Godfrey Harold Hardy published one of his most famous theorems. While seeking an elementary proof of the *Hilbert inequality*

$$\sum_{n,m=1}^{\infty} \frac{a_m a_n}{m+n} \leq \pi \sum_{n=1}^{\infty} a_n^2,$$

which holds for all non-negative square-summable sequences $a \in \ell^2(\mathbb{N})$, he unexpectedly uncovered a remarkable fact: the convergence of the series $\sum_{n=1}^{\infty} a_n^2$ implies the convergence of the series of squared cumulative averages $\sum_{n=1}^{\infty} (A_n/n)^2$, where $A_n := \sum_{k=1}^n a_k$. Now, for more than a century, this result has been widely recognized as the weak form of the *discrete Hardy inequality*.

Following Hardy's groundbreaking discovery, several leading mathematicians joined him in exploring and advancing this beautiful field of mathematics. Among the first was E. Landau, who, in 1921, extended the discrete Hardy inequality from sequences in ℓ^2 to those in ℓ^p for an arbitrary exponent $p > 1$. Namely, denoting $a_n := u_n - u_{n-1}$, for any finitely supported complex sequence $u \in C_0(\mathbb{N})$ with $u_0 := 0$, he established the *discrete p -Hardy inequality*

$$\sum_{n=1}^{\infty} |u_n - u_{n-1}|^p \geq \left(\frac{p-1}{p}\right)^p \sum_{n=1}^{\infty} \frac{|u_n|^p}{n^p} \quad (1)$$

along with the optimality of the constant on the right-hand side.

Shortly after, in 1924, Hardy formulated the well-known continuous form of the Hardy inequality, which reads

$$\int_0^{\infty} |\varphi'(x)|^p dx \geq \left(\frac{p-1}{p}\right)^p \int_0^{\infty} \frac{|\varphi(x)|^p}{x^p} dx \quad (2)$$

and holds for all smooth functions $\varphi \in C_0^{\infty}(\mathbb{R}_+)$ compactly supported in $\mathbb{R}_+ := (0, \infty)$. We refer the reader to [22, 23] for more historical details and proofs of inequalities (1) and (2).

Given the wide range of applications of the Hardy inequality in probability theory, the theory of partial differential equations, and numerous other domains, its various generalizations have been the subject of extensive investigation. Most of the research is focused on the linear case $p = 2$ where inequalities (1) and (2) represent lower bounds on the discrete and continuous Dirichlet Laplacians on the half-line in the sense of quadratic forms. In the context of the present work, we focus on higher-order variants of inequalities (1) and (2), which we refer to as the *Birman inequalities* or the *Hardy inequalities of higher order*. In 1961, Birman [3] generalized inequality (2), with $p = 2$, to higher-order derivatives:

$$\int_0^{\infty} |\varphi^{(\ell)}(x)|^2 dx \geq \left(\frac{1}{2}\right)_{\ell}^2 \int_0^{\infty} \frac{|\varphi(x)|^2}{x^{2\ell}} dx, \quad (3)$$

where $\varphi \in C_0^{\infty}(\mathbb{R}_+)$, $\ell \in \mathbb{N}$, and $(\nu)_{\ell} := \nu(\nu+1)\cdots(\nu+\ell-1)$ denotes the Pochhammer symbol for any $\nu \in \mathbb{R}$. A proof of (3), together with the optimality of the constant, can be found in [25]; see also [16] for other proofs and generalizations. The particular case $\ell = 2$ of (3) was discovered earlier by Rellich [26], even in higher dimensions, and is commonly referred to as the *Rellich inequality*.

Surprisingly, the discrete version of (3) was proven only recently in [18]. For any $\ell \in \mathbb{N}$ and $u \in C_0(\mathbb{N})$ with $u_n = 0$ for all $n < \ell$, the authors found the *discrete Birman inequality*

$$\sum_{n=\ell}^{\infty} |\nabla^{\ell} u_n|^2 \geq \left(\frac{1}{2}\right)_{\ell}^2 \sum_{n=\ell}^{\infty} \frac{|u_n|^2}{n^{2\ell}}, \quad (4)$$

where the *discrete gradient* is defined as $\nabla u_n := u_n - u_{n-1}$ for all $n \geq 1$ with the convention $u_0 := 0$; see Section 1 for details about the notation of discrete difference operators. The optimality of the constant on the right-hand side was proven therein.

The key idea behind their proof of the optimality relies on an interesting observation: assuming the discrete Hardy (or even Birman) inequality (1), one may derive its continuous version (2). Similarly, as E. Landau already noted in 1921, by examining a suitable step function, an integral version of the continuous Hardy inequality leads to the corresponding discrete inequality. This equivalence renders the following discrepancy of the two versions all the more intriguing.

It is a well-known fact that the classical Hardy weight $1/(4x^2)$ in the first-order continuous Hardy inequality

$$\int_0^{\infty} |\varphi'(x)|^2 dx \geq \frac{1}{4} \int_0^{\infty} \frac{|\varphi(x)|^2}{x^2} dx \quad (5)$$

is *critical*. This means that if (5) holds for all test functions $\varphi \in C_0^\infty(\mathbb{R}_+)$ with $1/(4x^2)$ replaced by a measurable function $\rho(x) \geq 1/(4x^2)$ for a.e. $x > 0$, then, necessarily, $\rho(x) = 1/(4x^2)$ for a.e. $x > 0$. In contrast, the opposite was found to be true in the discrete setting. In [19], M. Keller, Y. Pinchover, and F. Pogorzelski proved the discrete Hardy inequality

$$\sum_{n=1}^{\infty} |\nabla u_n|^2 \geq \sum_{n=1}^{\infty} \rho_n^{\text{KPP}} |u_n|^2, \quad (6)$$

for all $u \in C_0(\mathbb{N})$ with $u_0 := 0$, where the improved weight is given by the formula

$$\rho_n^{\text{KPP}} := 2 - \sqrt{\frac{n-1}{n}} - \sqrt{\frac{n+1}{n}} > \frac{1}{4n^2}$$

for all $n \in \mathbb{N}$. Additionally, the authors proved in [20] that the weight ρ^{KPP} is *null-critical* (see Remark 2.2 for the definition), which, in particular, implies that the weight is *critical*, i.e. that it cannot be further improved by a pointwise bigger sequence. This result was subsequently extended to an arbitrary exponent $p > 1$ within the dissertation of F. Fisher [11], namely, in the article [13] coauthored with M. Keller and F. Pogorzelski, where the p -Hardy inequality

$$\sum_{n=1}^{\infty} |u_n - u_{n-1}|^p \geq \sum_{n=1}^{\infty} \rho_n^{\text{FKP}} |u_n|^p, \quad (7)$$

with the null-critical weight

$$\rho_n^{\text{FKP}} := \left(1 - \left(1 - \frac{1}{n}\right)^{1-1/p}\right)^{p-1} - \left(\left(1 + \frac{1}{n}\right)^{1-1/p} - 1\right)^{p-1} > \left(\frac{p-1}{p}\right)^p \frac{1}{n^p} \quad (8)$$

was verified for all real-valued $u \in C_0(\mathbb{N})$ with $u_0 := 0$.

Naturally, it was expected that even the higher-order discrete Hardy inequalities (4) admit an improvement analogous to (6). Initially, some refinements were made in the Rellich case $\ell = 2$ in [18, 15], and later the hypothesis was confirmed in full generality in [32], where we established the *optimal* discrete Hardy inequality of an arbitrary order $\ell \in \mathbb{N}$:

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^2 \geq \sum_{n=\ell}^{\infty} \rho_n^{(\ell)} |u_n|^2$$

for all $u \in C_0(\mathbb{N})$ with $u_n = 0$ for all $n < \ell$, where the weights are given by the formulas

$$\rho_n^{(\ell)} := \frac{(-\Delta)^\ell \mathfrak{g}_n^{(\ell)}}{\mathfrak{g}_n^{(\ell)}} \quad \text{for all } n \geq \ell \quad \text{and} \quad \mathfrak{g}_n^{(\ell)} := \sqrt{n} \prod_{j=1}^{\ell-1} (n-j) \quad \text{for all } n \geq 0. \quad (9)$$

Here the *discrete Laplacian* is defined as $\Delta u_n := u_{n-1} - 2u_n + u_{n+1}$; confer also with Section 1. *Optimality* is a rather strong property (see Definition 2.1), which, in particular, entails null-criticality and hence criticality of the Hardy weight under consideration.

While both the discrete and continuous Birman inequalities are well-established classical results for the case $p = 2$, the nonlinear setting $p \neq 2$ is substantially less developed in the existing literature. Although the continuous p -Hardy inequality of order $\ell \in \mathbb{N}$,

$$\int_0^\infty |\varphi^{(\ell)}(x)|^p dx \geq \left(\frac{p-1}{p}\right)_\ell^p \int_0^\infty \frac{|\varphi(x)|^p}{x^{\ell p}} dx, \quad (10)$$

which holds for all $p > 1$ and all $\varphi \in C_0^\infty(\mathbb{R}_+)$, is probably well-known among experts, we have been unable to identify a reference in which this statement is formulated explicitly, with the exception of the particular case $\ell = 2$, corresponding to the p -Rellich inequality, which appears in [25, Eq. (1.2)]. At an implicit level, inequality (10) can be traced back to the classical work of Hardy, Littlewood, and Pólya [17]; see Remark 4.12. Related results in higher dimensions can also be found in [8].

Furthermore, an essential component of the framework is currently missing, leaving the theory of Hardy inequalities on the half-line incomplete: the discrete p -Hardy inequalities of higher order. The primary objective of this paper is to address this gap by establishing and further analyzing new explicit discrete p -Birman inequalities.

This paper is structured as follows:

- In Section 1, we introduce the necessary notation and compare it with the conventions adopted in recent works [21, 18, 32] on Hardy-type inequalities.

- Building on the approach developed in [32], where the analysis was restricted to the more tractable linear case $p = 2$, we generalize these results to the nonlinear setting $p > 1$ in Section 2. Through Theorems 2.7, 2.9, and 2.10, we develop a method to construct abstract p -Birman weights in terms of a parameter sequence $\mathbf{g}$ subject to suitable monotonicity and growth conditions.
- In Section 3, we apply the results of Section 2 to the parameter sequences $\mathbf{g}^{(\ell,p)}$ and $\tilde{\mathbf{g}}^{(\ell,p)}$, as defined in (2.8) and (2.9) respectively, and obtain explicit optimal p -Hardy weights of an arbitrary order.
- In Section 4, we establish classical discrete p -Birman inequalities, which constitute direct discrete analogs of the continuous inequality (10) and show that the optimal weights $\rho^{(\ell,p)}$, introduced in Section 3, improve upon the corresponding classical weights. Finally, in Section 4.3, we present a new proof of the continuous result by recovering it as a limiting case of its discrete analog.

1. NOTATION OF DISCRETE DIFFERENCE OPERATORS

The notation of discrete difference operators used in papers studying discrete Hardy inequalities [21, 18, 32] varies and therefore may cause confusion for a non-specialized reader. In this section, our aim is to clarify, but mainly unify, the miscellaneous conventions.

First, let us consider the Lebesgue space of square-summable sequences $\ell^2(\mathbb{N})$ indexed by natural numbers $\mathbb{N} := \{1, 2, \dots\}$, and its standard basis $\{\delta_n\}_{n=1}^\infty$ defined as

$$(\delta_n)_m := \delta_{nm} = \begin{cases} 0 & \text{if } n \neq m, \\ 1 & \text{if } n = m, \end{cases} \quad \text{for all } m, n \in \mathbb{N},$$

where δ is the Kronecker delta. On this space, we define the *discrete gradient* and *divergence* by the formulas

$$\nabla u_n := \begin{cases} u_n - u_{n-1} & \text{if } n \geq 2, \\ u_1 & \text{if } n = 1, \end{cases} \quad \text{and} \quad \operatorname{div} u_n := u_{n+1} - u_n \quad \text{for all } n \in \mathbb{N}.$$

Furthermore, the *discrete Laplacian* on $\ell^2(\mathbb{N})$ is then given by the familiar formula $\Delta_{\mathbb{N}} := \operatorname{div} \nabla$, i.e.

$$(\Delta_{\mathbb{N}} u)_n := \begin{cases} u_{n-1} - 2u_n + u_{n+1} & \text{if } n \geq 2, \\ -2u_1 + u_2 & \text{if } n = 1. \end{cases}$$

The index $\mathbb{N}$ is used to distinguish the operator from the Laplace operator used in the rest of the text, defined below by the formula (1.4). The operator $-\Delta_{\mathbb{N}}$ also coincides with the *graph Laplacian* associated with the standard graph structure on $\mathbb{N}$, which was studied in [21]. Notice that on the Hilbert space $\ell^2(\mathbb{N})$ endowed with the Euclidean inner product, the adjoint of the discrete gradient is $\nabla^* = -\operatorname{div}$, hence $-\Delta_{\mathbb{N}} = \nabla^* \nabla$ is a positive self-adjoint operator (we shall omit $\circ$ when composing difference operators for concise notation).

In [18], the authors defined the half-integer powers of the Laplace operator as

$$(-\Delta_{\mathbb{N}})^{\ell/2} := \begin{cases} (-\Delta_{\mathbb{N}})^m & \text{if } \ell = 2m \in 2\mathbb{N}, \\ \nabla(-\Delta_{\mathbb{N}})^m & \text{if } \ell = 2m + 1 \in 2\mathbb{N}_0 + 1, \end{cases}$$

and considered Hardy inequalities of higher order (4) in the form

$$\sum_{n=\lceil \ell/2 \rceil}^{\infty} |(-\Delta_{\mathbb{N}})^{\ell/2} u_n|^2 \geq \left(\frac{1}{2}\right)^2 \sum_{n=\ell}^{\infty} \frac{|u_n|^2}{n^{2\ell}}, \quad (1.1)$$

for all $u \in \ell^2(\mathbb{N})$ satisfying $u_n = 0$ for all $n < \ell$, where $\lceil \nu \rceil$ stands for the smallest integer greater than or equal to $\nu \in \mathbb{R}$. With this convention, the left-hand side can be rewritten as the quadratic form of the operator $(-\Delta_{\mathbb{N}})^\ell$ since

$$\sum_{n=\lceil \ell/2 \rceil}^{\infty} |(-\Delta_{\mathbb{N}})^{\ell/2} u_n|^2 = \|(-\Delta_{\mathbb{N}})^{\ell/2} u\|^2 = \langle u, ((-\Delta_{\mathbb{N}})^{\ell/2})^* (-\Delta_{\mathbb{N}})^{\ell/2} u \rangle = \langle u, (-\Delta_{\mathbb{N}})^\ell u \rangle.$$

Therefore, the inequality (1.1) amounts to the lower bound

$$(-\Delta_{\mathbb{N}})^\ell \geq \rho^B \quad \text{on the subspace } \{\delta_1, \dots, \delta_{\ell-1}\}^\perp \subset \ell^2(\mathbb{N}), \quad (1.2)$$

where ρ^B stands for the operator of multiplication by the classical Birman weight ρ^B given by the right-hand side of (1.1).

On the other hand, we will use the notation that utilizes solely the discrete gradient, where the left-hand side of the Birman inequality (4) reads

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^2 = \|\nabla^\ell u\|^2 = \langle u, (-1)^\ell \operatorname{div}^\ell \nabla^\ell u \rangle.$$

The Birman inequality then takes the form

$$(-1)^\ell \operatorname{div}^\ell \nabla^\ell \geq \rho^B \quad \text{on the subspace } \{\delta_1, \dots, \delta_{\ell-1}\}^\perp \subset \ell^2(\mathbb{N}). \quad (1.3)$$

It is important to emphasize the subtle differences between the operators $(-\Delta_{\mathbb{N}})^\ell$ and $(-1)^\ell \operatorname{div}^\ell \nabla^\ell$. While the matrix elements of $(-1)^\ell \operatorname{div}^\ell \nabla^\ell$ with respect to the standard basis can be straightforwardly computed as

$$\left((-1)^\ell \operatorname{div}^\ell \nabla^\ell \right)_{mn} = (-1)^{m+n} \binom{2\ell}{\ell+m-n} \quad \text{for all } m, n \in \mathbb{N},$$

and define a symmetric banded Toeplitz (i.e. diagonal-constant) matrix on $\ell^2(\mathbb{N})$, the matrix elements of $(-\Delta_{\mathbb{N}})^\ell$ are given by [14, Proposition 2.1] as

$$(-\Delta_{\mathbb{N}})_{mn}^\ell = (-1)^{m+n} \left[\binom{2\ell}{\ell+m-n} - \binom{2\ell}{\ell+m+n} \right] \quad \text{for all } m, n \in \mathbb{N}.$$

Observe that the operator $(-\Delta_{\mathbb{N}})^\ell$ is no longer Toeplitz and differs from $(-1)^\ell \operatorname{div}^\ell \nabla^\ell$ by an upper-left $(\ell-1) \times (\ell-1)$ Hankel matrix (i.e. a matrix with constant anti-diagonals). Nevertheless, the operators coincide on the subspace $\{\delta_1, \dots, \delta_{\ell-1}\}^\perp \subset \ell^2(\mathbb{N})$, thus rendering the inequalities (1.2) and (1.3) equivalent.

The operator $(-1)^\ell \operatorname{div}^\ell \nabla^\ell$ can also be interpreted as the restriction of the ℓ -th power of the discrete Laplacian defined on the Hilbert space $\ell^2(\mathbb{Z})$ by

$$(\Delta u)_n := u_{n-1} - 2u_n + u_{n+1} \quad \text{for all } n \in \mathbb{Z}. \quad (1.4)$$

Indeed, direct computation yields the corresponding matrix representation with respect to the canonical basis $\{\delta_n\}_{n \in \mathbb{Z}}$ of $\ell^2(\mathbb{Z})$ in the form

$$(-\Delta)_{mn}^\ell = (-1)^{m+n} \binom{2\ell}{\ell+m-n} \quad \text{for all } m, n \in \mathbb{Z},$$

hence $(-\Delta)^\ell|_{\ell^2(\mathbb{N})} = (-1)^\ell \operatorname{div}^\ell \nabla^\ell$.

Keeping these considerations in mind, we formally extend the sequences in question by zero for negative indices and introduce the notation employed throughout this paper, which is adopted from our work [32]. We denote by $C(\mathbb{Z})$ the space of all complex-valued sequences indexed by $\mathbb{Z}$ and, for each $\ell \in \mathbb{N}$, we define the subspace

$$H^\ell := \{u \in C(\mathbb{Z}) \mid u_n = 0 \text{ for all } n < \ell\}$$

and its subspaces

$$\mathcal{H}^{\ell,p} := H^\ell \cap \ell^p(\mathbb{Z}) \quad \text{and} \quad \mathcal{H}_0^\ell := H^\ell \cap C_0(\mathbb{Z}),$$

endowed with the ℓ^p -norm $\|\cdot\|_p$, where $C_0(\mathbb{Z})$ denotes the subspace of $C(\mathbb{Z})$ consisting of all finitely supported sequences. With a slight abuse of notation, we additionally introduce the *discrete gradient* and *discrete divergence* operators acting on $C(\mathbb{Z})$ as

$$(\nabla u)_n := u_n - u_{n-1} \quad \text{and} \quad (\operatorname{div} u)_n := u_{n+1} - u_n \quad (1.5)$$

for all $n \in \mathbb{Z}$. Since these operators commute on $C(\mathbb{Z})$, this convention simplifies subsequent arguments, as it obviates the need to specify the subspaces on which the identity $\operatorname{div} \nabla = \nabla \operatorname{div} = \Delta$ holds. We shall also make use of the *forward shift operator* S and the *middle operator* M defined on $C(\mathbb{Z})$ as

$$Su_n := u_{n+1} \quad \text{and} \quad Mu_n := \frac{u_{n+1} + u_n}{2} \quad (1.6)$$

for all $n \in \mathbb{Z}$. Note also that we have the operator identities $S^{-1}S = SS^{-1} = I$, $\operatorname{div} = S - I$, $\nabla = I - S^{-1}$, and therefore $\operatorname{div} = \nabla S = S \nabla$.

For $p \neq 2$, the absence of an inner product on $\ell^p(\mathbb{Z})$ makes it necessary to introduce further notation. Although the p -Birman inequalities cannot be interpreted in the sense of quadratic forms, it is nevertheless possible to *polarize* the ℓ^p -norm of $\nabla^\ell u$, for $u \in \mathcal{H}_0^\ell$, as

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p = \sum_{n=\ell}^{\infty} \overline{u_n} (-\Delta_p^{(\ell)} u)_n,$$

where the nonlinear p -Laplace operator of order $\ell \in \mathbb{N}$ is defined by

$$-\Delta_p^{(\ell)} := (-1)^\ell \operatorname{div}^\ell (\nabla^\ell \cdot)^{(p-1)} \quad \text{and} \quad -\Delta_p^{(0)} := (\cdot)^{(p-1)}. \quad (1.7)$$

Moreover, for streamlined notation, we set $\nu^{(\lambda)} := \nu |\nu|^{\lambda-1}$ for all $\nu \in \mathbb{C}$ and $\lambda > 0$, with the convention $0^{(\lambda)} = 0$.

2. ABSTRACT OPTIMAL p -BIRMAN INEQUALITIES

In this section, we explain how to construct p -Hardy weights of any order $\ell \in \mathbb{N}$ given a *parameter sequence* $\mathbf{g}$ that satisfies certain sufficient conditions. In fact, we slightly modify and extend the strategy from [32], where we considered the linear case $p = 2$, to the more involved general case with an arbitrary $p > 1$. Furthermore, we will discuss optimality of the acquired weights under stronger assumptions on the asymptotic behavior of $\mathbf{g}_n$ as $n \rightarrow \infty$.

We begin by presenting the main results of this section. To avoid distracting the reader and to maintain clarity, the corresponding proofs are postponed to Section 2.1. First, let us recall the definition of p -Hardy weights of higher order and their optimality; see [12] and [32].

Definition 2.1. Let $\ell \in \mathbb{N}$ and $p > 1$. A non-negative sequence $\{\rho_n\}_{n=\ell}^\infty$ is called a *discrete p -Hardy weight of order ℓ* (or *discrete p -Birman weight*) if and only if the *discrete p -Hardy inequality of order ℓ* (or *discrete p -Birman inequality*)

$$\sum_{n=\ell}^\infty |\nabla^\ell u_n|^p \geq \sum_{n=\ell}^\infty \rho_n |u_n|^p \quad (2.1)$$

holds for any $u \in \mathcal{H}_0^\ell$. In addition, the weight ρ is said to be *optimal* if it exhibits the following three properties:

- (i) *Criticality*: The weight ρ is called *critical* if for any p -Hardy weight $\tilde{\rho}$ of order ℓ , such that $\tilde{\rho}_n \geq \rho_n$ for all $n \geq \ell$, it follows that $\tilde{\rho} = \rho$.
- (ii) *Non-attainability*: The weight ρ is called *non-attainable* if, whenever the equality in (2.1) is attained for $u \in \mathcal{H}_0^\ell$ such that the right-hand side is finite, necessarily $u \equiv 0$.
- (iii) *Optimality near infinity*: The weight ρ is called *optimal near infinity* if for any $M \geq \ell$ and $\varepsilon > 0$ there exists $u \in \mathcal{H}_0^M$ such that

$$\sum_{n=\ell}^\infty |\nabla^\ell u_n|^p < (1 + \varepsilon) \sum_{n=\ell}^\infty \rho_n |u_n|^p.$$

Remark 2.2. We complement the three optimality properties with several explanatory remarks:

- (i) Criticality means that the p -Hardy weight enjoying this property cannot be further improved by a pointwise bigger sequence.
- (ii) The notion of non-attainability is not uniform in the existing literature. In [20], where first-order Hardy inequalities on graphs are investigated, the property that corresponds to the combination of criticality and non-attainability is referred to as *null-criticality*. Using the criticality theory for Schrödinger operators on discrete graphs, as developed in [21], the authors analyze whether the unique positive *generalized eigenvector* of the operator $-\Delta - \rho$, called the *Agmon ground state*, is, in fact, a genuine eigenvector. This definition was later extended to the ℓ^p setting in [12, 10]. Moreover, it was shown in [11, Theorem 11.2] that for certain p -Hardy weights, null-criticality also implies optimality near infinity. Nevertheless, it is evident that the notions of criticality and non-attainability employed in this work actually entail null-criticality and therefore constitute a stronger condition.

We additionally introduce a slightly weaker property, which we refer to as *non-attainability in $\mathcal{H}_0^\ell$* . Specifically, the weight ρ is said to be *non-attainable in $\mathcal{H}_0^\ell$* if, whenever equality in (2.1) is attained for $u \in \mathcal{H}_0^\ell$, necessarily $u \equiv 0$.

- (iii) Optimality near infinity means that the constant 1 appearing on the right-hand side of (2.1) cannot be further improved, even if we restrict ourselves only to the sequences from the space $\mathcal{H}_0^M \subset \mathcal{H}_0^\ell$ for any $M \geq \ell$. It is equivalent to the identity

$$\inf_{u \in \mathcal{H}_0^M \setminus \{0\}} \frac{\sum_{n=\ell}^\infty |\nabla^\ell u_n|^p}{\sum_{n=\ell}^\infty \rho_n |u_n|^p} = 1 \quad \text{for all } M \geq \ell. \quad (2.2)$$

Let us also remark that optimality does not characterize the weight uniquely as was observed in [32], where we found infinitely many optimal Hardy weights of any order ℓ for $p = 2$.

Our starting point shall be a weighted generalization of an abstract first-order p -Hardy inequality analogous to [10, Theorem 3.1]. In [10], the authors study p -Hardy inequalities in an abstract graph setting for real-valued sequences. For our purposes, we consider their results for the standard line graph on $\mathbb{N}_0$. The proof of the weighted generalization is reminiscent of the one in [10] or [13, Proposition 4], but for the reader's convenience and completeness, it will be revisited in Section 2.1 nonetheless. The inequality is potentially of independent interest and will be employed to derive the Copson inequality (4.12) in Section 4.

Unless explicitly stated otherwise, we shall always assume $\ell \in \mathbb{N}$ and $p, q > 1$ to be Hölder conjugate exponents, i.e., they satisfy $1/p + 1/q = 1$, throughout the following. Moreover, for two real numbers (or functions) α, β , we write $\alpha \lesssim \beta$ if and only if there exists a positive constant $C_p > 0$ such that $\alpha \leq C_p \beta$, where the constant C_p may depend only on p . We write $\alpha \approx \beta$ if and only if $\alpha \lesssim \beta \lesssim \alpha$.

Theorem 2.3. *Let $V \in C(\mathbb{N})$, such that $V_n \geq 0$ for all $n \geq 1$, and $\mathbf{g} \in H^1$, such that $\mathbf{g}_{n+1} \geq \mathbf{g}_n > 0$ for all $n \geq 1$. Then for any $u \in \mathcal{H}_0^1$, we have the inequalities*

$$\sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(q; \mathbf{g}, u)_n \lesssim \sum_{n=1}^{\infty} V_n |\nabla u_n|^p + \sum_{n=1}^{\infty} \frac{\operatorname{div}(V(\nabla \mathbf{g})^{p-1})_n}{\mathbf{g}_n^{p-1}} |u_n|^p \lesssim \sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(p; \mathbf{g}, u)_n \quad (2.3)$$

where the remainder terms are given by the formulas

$$\mathcal{R}(s; \mathbf{g}, u)_n := \begin{cases} \left| \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n+1}}} u_{n+1} - \sqrt{\frac{\mathbf{g}_{n+1}}{\mathbf{g}_n}} u_n \right|^p & \text{if } s \in (1, 2], \\ \left| \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n+1}}} u_{n+1} - \sqrt{\frac{\mathbf{g}_{n+1}}{\mathbf{g}_n}} u_n \right|^2 \left(|\operatorname{div} u_n| + \frac{|u_n|}{\mathbf{g}_n} \operatorname{div} \mathbf{g}_n \right)^{p-2} & \text{if } s \in (2, \infty), \end{cases} \quad (2.4)$$

for all $n \geq 1$, with the convention $0^{p-2} \equiv 0$ if $p \in (1, 2]$.

Remark 2.4. The parameter s appearing in the remainder terms, defined by (2.4), serves as an indicator specifying which of the two formulas is to be used for the lower and the upper bound in (2.3). Since the Hölder conjugate exponents $p, q > 1$ satisfy $p > 2$ if and only if $q \in (1, 2)$, this notation encodes the fact that the two distinct remainder terms occurring on the left-hand and right-hand sides of the inequality (2.3) are interchanged when passing from the regime $p \in (1, 2)$ to the regime $p > 2$. The lower estimate in (2.3) asserts an abstract weighted p -Hardy inequality for any complex-valued sequence $u \in \mathcal{H}_0^1$. On the other hand, the upper estimate in (2.3) indicates how tight the inequality is, which will be essential later to derive the optimality of particular p -Hardy weights. Both estimates are rooted in the fundamental inequalities given by Lemma 2.13.

Remark 2.5. In the case $p = 2$, the inequalities in (2.3) hold as equalities with the implicit constants in the estimates $\lesssim$ being equal to 1. This is analogous to the results of [32, Theorem 0], which asserts that for any $V \in C(\mathbb{N})$ and $\mathbf{g} \in H^1$, such that $\mathbf{g}_n > 0$ for all $n \geq 1$, we have the identity

$$\sum_{n=1}^{\infty} V_n |\nabla u_n|^2 + \sum_{n=1}^{\infty} \frac{\operatorname{div}(V \nabla \mathbf{g})_n}{\mathbf{g}_n} |u_n|^2 = \sum_{n=1}^{\infty} V_{n+1} \left| \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n+1}}} u_{n+1} - \sqrt{\frac{\mathbf{g}_{n+1}}{\mathbf{g}_n}} u_n \right|^2$$

for all $u \in \mathcal{H}_0^1$. Note that the parameter sequence $\mathbf{g}$ is no longer required to be non-decreasing, nor is the potential V required to be non-negative.

Remark 2.6. In [10], the authors restrict their analysis to real-valued sequences, which makes the situation less technically involved. In particular, if $V \in C(\mathbb{N})$, such that $V_n \geq 0$ for all $n \geq 1$, and $\mathbf{g} \in H^1$, such that $\mathbf{g}_n > 0$ for all $n \geq 1$ ($\mathbf{g}$ no longer has to be non-decreasing), we have the equivalence

$$\sum_{n=1}^{\infty} V_n |\nabla u_n|^p + \sum_{n=1}^{\infty} \frac{\operatorname{div}(V(\nabla \mathbf{g})^{(p-1)})_n}{\mathbf{g}_n^{p-1}} |u_n|^p \approx \sum_{n=1}^{\infty} V_{n+1} \tilde{\mathcal{R}}(p; \mathbf{g}, u)_n,$$

for all real-valued sequences $u \in \mathcal{H}_0^1$, where the remainder is given by the formula

$$\tilde{\mathcal{R}}(p; \mathbf{g}, u)_n := \left| \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n+1}}} u_{n+1} - \sqrt{\frac{\mathbf{g}_{n+1}}{\mathbf{g}_n}} u_n \right|^2 \left(\left| \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n+1}}} u_{n+1} - \sqrt{\frac{\mathbf{g}_{n+1}}{\mathbf{g}_n}} u_n \right| + M \left(\frac{|u_n|}{\mathbf{g}_n} \right) |\operatorname{div} \mathbf{g}_n| \right)^{p-2} \quad (2.5)$$

for all $p > 1$ and $n \geq 1$. The proof is analogous to the proof of Theorem 2.3 and is therefore omitted; see [10, Lemma 3.8, Eq. (3.9)] for the corresponding underlying inequality.

The weight V in inequality (2.3) enables an iterative procedure that yields analogous inequalities for higher-order difference operators. The key idea behind the derivation of such inequalities is a suitable selection of the parameter sequence $\mathbf{g}$ in each step of the iteration. In this way, we obtain convenient representations of the weights V in terms of $\mathbf{g}$, which, in turn, allows us to guarantee their non-negativity under additional assumptions imposed on $\mathbf{g}$.

Theorem 2.7. Suppose that $\mathbf{g} \in H^\ell$ satisfies the assumptions

$$\nabla^k \mathbf{g}_n > 0 \quad \text{for all } n \geq \ell \text{ and } k \in \{0, \dots, \ell\}, \quad (\text{A1})$$

$$-\Delta_p^{(\ell-k)} \nabla^k \mathbf{g}_n \geq 0 \quad \text{for all } n \geq \ell \text{ and } k \in \{1, \dots, \ell-1\}. \quad (\text{A2})$$

Then for all $u \in \mathcal{H}_0^\ell$, we have the inequalities

$$\sum_{k=1}^{\ell} R_k^{(\ell)}(q; \mathbf{g}, u) \lesssim \sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p - \sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell)} \mathbf{g}_n}{\mathbf{g}_n^{p-1}} |u_n|^p \lesssim \sum_{k=1}^{\ell} R_k^{(\ell)}(p; \mathbf{g}, u), \quad (\text{2.6})$$

where the remainders are given by the formula

$$R_k^{(\ell)}(s; \mathbf{g}, u) := \sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell-k)} \nabla^k \mathbf{g}_{n+1}}{(\nabla^k \mathbf{g}_{n+1})^{p-1}} \mathcal{R}(s; \nabla^{k-1} \mathbf{g}, \nabla^{k-1} u)_n \quad (\text{2.7})$$

with $\mathcal{R}(s; \mathbf{g}, u)$ defined by (2.4).

Remark 2.8. In view of Remark 2.6, for real-valued sequences, we may deduce an equivalence analogous to inequalities (2.6) with remainders corresponding to (2.5). Namely, for $\mathbf{g} \in H^\ell$ satisfying assumptions (A1) and (A2), we have the equivalence

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p - \sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell)} \mathbf{g}_n}{\mathbf{g}_n^{p-1}} |u_n|^p \approx \sum_{k=1}^{\ell} \left[\sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell-k)} \nabla^k \mathbf{g}_{n+1}}{(\nabla^k \mathbf{g}_{n+1})^{p-1}} \tilde{\mathcal{R}}(p; \nabla^{k-1} \mathbf{g}, \nabla^{k-1} u)_n \right],$$

for all real-valued sequences $u \in \mathcal{H}_0^\ell$. The proof works along the lines of the proof of Theorem 2.7 in Section 2.1.2, hence it is omitted as well.

Let us further assume that the inequality in assumption (A2) holds for $k = 0$, that is, $-\Delta_p^{(\ell)} \mathbf{g}_n \geq 0$ for all $n \geq \ell$. Then

$$\rho(\mathbf{g}) := \frac{-\Delta_p^{(\ell)} \mathbf{g}}{\mathbf{g}^{p-1}}$$

defines a non-negative sequence in H^ℓ , and thus constitutes a p -Hardy weight of order ℓ . Moreover, we prove that this weight is not attainable in $\mathcal{H}_0^\ell$; see Remark 2.2 for the definition.

Theorem 2.9. Suppose that $\mathbf{g} \in H^\ell$ satisfies assumptions (A1), (A2), and, in addition,

$$-\Delta_p^{(\ell)} \mathbf{g}_n \geq 0 \quad \text{for all } n \geq \ell. \quad (\text{A3})$$

Then $\rho(\mathbf{g}) = -\Delta_p^{(\ell)} \mathbf{g} / \mathbf{g}^{p-1}$ is a discrete p -Hardy weight of order ℓ , which is non-attainable in $\mathcal{H}_0^\ell$.

By imposing additional requirements on the asymptotic behavior of $\mathbf{g}_n$, as $n \rightarrow \infty$, and strict positivity in assumption (A3), we obtain sufficient conditions for the optimality of the discrete p -Hardy weights $\rho(\mathbf{g})$. We consider the following statement to be our first main result.

Theorem 2.10. Suppose that $\mathbf{g} \in H^\ell$ satisfies assumptions (A1), (A2), and (A3) with strict positivity, i.e. $-\Delta_p^{(\ell)} \mathbf{g}_n > 0$ for all $n \geq \ell$. In addition, suppose that $\mathbf{g}$ admits the asymptotic expansion

$$\mathbf{g}_n = \sum_{j=0}^{2\ell} \alpha_j n^{\ell-j-1/p} + \mathcal{O}(n^{-\ell-1-1/p}), \quad \text{as } n \rightarrow \infty, \text{ for some } \alpha_j \in \mathbb{R} \text{ with } \alpha_0 \neq 0. \quad (\text{A4})$$

Then the discrete p -Hardy weight $\rho(\mathbf{g}) = -\Delta_p^{(\ell)} \mathbf{g} / \mathbf{g}^{p-1}$ of order ℓ is optimal.

Remark 2.11. The strict positivity in assumption (A3) is required solely for the non-attainability of the weight $\rho(\mathbf{g})$, whereas the properties of criticality and optimality near infinity are already ensured by assumptions (A1)–(A4).

Remark 2.12. Without going into technical details, we note that the asymptotic expansion in (A4) can be replaced by one depending on a parameter $s \in (0, 1)$. The admissible range of s then determines which of the properties from the definition of optimality 2.1 are satisfied by the corresponding weight $\rho(\mathbf{g})$. Specifically, for $\mathbf{g} \in H^\ell$ subject to assumptions (A1), (A2), (A3) with strict positivity, and, in addition,

$$\mathbf{g}_n = \sum_{j=0}^{2\ell} \alpha_j n^{\ell-j-s} + \mathcal{O}(n^{-\ell-1-s}) \quad \text{as } n \rightarrow \infty, \text{ for some } \alpha_j \in \mathbb{R} \text{ with } \alpha_0 \neq 0, \quad (\text{A4}')$$

the discrete p -Hardy weight $\rho(\mathbf{g})$ of order ℓ is critical if $s \in [1/p, 1)$, non-attainable if $s \in (0, 1/p]$, and optimal near infinity if $s = 1/p$. Since the proof follows almost verbatim the arguments in Section 2.1.4, it will be omitted; see also [32, Theorem 3] for an analogous result in the linear case $p = 2$.

The final piece of the puzzle is to construct an explicit parameter sequence $\mathbf{g} \in H^\ell$ that fulfills all the assumptions (A1)–(A4). To this end, in Section 3, we shall examine the following candidate sequences:

- First, we consider a parameter sequence constructed in the form of a suitable Gamma function ratio:

$$\mathbf{g}_n^{(\ell,p)} := \frac{\Gamma(n + 1/q)}{\Gamma(n - \ell + 1)} \quad \text{for all } n \in \mathbb{Z}, \quad (2.8)$$

where Γ is the *Euler Gamma function*, and we consider $1/\Gamma$ to be an entire function with zeros at the points $0, -1, \dots$. In Section 3.1, we show that $\mathbf{g}^{(\ell,p)}$ gives rise to an optimal p -Hardy weight of order ℓ .

In particular, for $\ell = 1$, we discover the optimal first-order p -Hardy weight

$$\rho_n^{(1,p)} := \left(\frac{p-1}{p}\right)^{p-1} \left[\frac{1}{(n-1/p)^{p-1}} - \frac{1}{n^{p-1}} \right] \quad \text{for all } n \geq 1,$$

while in the linear case $p = 2$, we obtain the optimal Hardy weights of order ℓ in the form

$$\rho_n^{(\ell,2)} := \left(\frac{1}{2}\right)_\ell^2 \frac{1}{n(n-1/2)(n-1)\cdots(n-\ell+1/2)} \quad \text{for all } n \geq \ell.$$

These results complement the earlier works [13, 32], in which the optimal Hardy inequalities with the weights ρ^{FKP} and $\rho^{(\ell)}$, given by formulas (8) and (9), respectively, were established, and yield new weight sequences of a simpler and more explicit form. On the other hand, prior to this work, no (optimal) p -Hardy weights of order $\ell \geq 2$ had been identified.

- Second, in light of the choice of $\mathbf{g}$ for $p = 2$ in (9), as well as the first-order p -Hardy weight (8) introduced by Fisher–Keller–Pogorzelski [13], it is natural to investigate also the following extension:

$$\tilde{\mathbf{g}}_n^{(\ell,p)} := n^{1-1/p} \prod_{j=1}^{\ell-1} (n-j) \quad \text{for all } n \geq 0 \quad \text{and} \quad \tilde{\mathbf{g}}_n := 0 \quad \text{for all } n < 0. \quad (2.9)$$

Even though it is fairly easy to verify assumptions (A1) and (A4) for the sequence (2.9), the nonlinearity of the discrete p -Laplace operator $-\Delta_p^{(k)}$ of order $k \in \mathbb{N}$ makes validating assumptions (A2) and (A3) a very non-trivial problem. Regrettably, we were unable to verify these assumptions in a general setting given by a combination of $p > 1$ and $\ell \in \mathbb{N}$. Partial results are summarized in Section 3.2.

2.1. Proofs. In this section, we systematically derive the proofs of Theorems 2.3, 2.7, 2.9, and 2.10.

2.1.1. Proof of Theorem 2.3. The proof of the inequalities in (2.3) relies on the following lemma, which provides auxiliary fundamental estimates and extends [10, Lemma 3.8] to the complex domain.

Lemma 2.13. *Let $t \in [0, 1]$. Then for any complex number $z \in \mathbb{C}$, we have the inequalities*

$$t|z-1|^2(|z-t|+1-t)^{p-2} \lesssim |z-t|^p - (1-t)^{p-1}(|z|^p - t) \lesssim t^{p/2}|z-1|^p \quad \text{if } p \in (1, 2], \quad (2.10)$$

$$t|z-1|^2(|z-t|+1-t)^{p-2} \gtrsim |z-t|^p - (1-t)^{p-1}(|z|^p - t) \gtrsim t^{p/2}|z-1|^p \quad \text{if } p \in (2, \infty). \quad (2.11)$$

(The left-hand side of (2.10) is to be understood as zero if $z = t = 1$.)

Proof (of Lemma 2.13). The proof is decomposed into two parts, in which we establish

- a) the lower bound for $p \in (1, 2]$, and, respectively, the upper bound for $p \in (2, \infty)$,
- b) the upper bound for $p \in (1, 2]$, and, respectively, the lower bound for $p \in (2, \infty)$,

for the term $|z-t|^p - (1-t)^{p-1}(|z|^p - t)$. In both cases, the strategy is to extend the corresponding real-variable estimates, originally formulated in [10, Lemma 3.8], to the complex setting. However, the methods employed in a) and b) differ, which requires treating the two parts of the proof separately.

a) Suppose $p \in (1, 2]$. By [10, Lemma 3.8, Eq. (3.8)], there exists $C_p > 0$ sufficiently small, such that the inequality

$$C_p t |x-1|^2 (|x-t|+1-t)^{p-2} \leq |x-t|^p - (1-t)^{p-1}(|x|^p - t) \quad (2.12)$$

holds for all $t \in [0, 1]$ and $x \in \mathbb{R}$. For $z \in \mathbb{C}$, write $z = re^{i\theta}$, where $r = |z| \geq 0$ and $\theta \in [0, 2\pi)$ is the phase of z . Next, for $u := \cos \theta \in [-1, 1]$, denote

$$A(u) := |z-t| = \sqrt{r^2 - 2rtu + t^2} \quad \text{and} \quad B(u) := |z-1| = \sqrt{r^2 - 2ru + 1}.$$

We will show, that the auxiliary function

$$F(u) := A^p(u) - (1-t)^{p-1}(r^p - t) - C_p t B^2(u)(A(u) + 1-t)^{p-2},$$

defined on the interval $[-1, 1]$, is non-increasing, and hence attains its minimum at $u = 1$. Since $F(1) \geq 0$ by inequality (2.12), it follows that $F(u) \geq 0$ for all $u \in [-1, 1]$, which amounts to the lower estimate in (2.10). Indeed, a routine calculation yields

$$\begin{aligned} \frac{dF}{du} &= -prtA^{p-2} + C_p rt \left[2(A+1-t)^{p-2} + (p-2)t \frac{B^2}{A}(A+1-t)^{p-3} \right] \\ &\leq rt [2C_p(A+1-t)^{p-2} - pA^{p-2}] \leq rtA^{p-2}(2C_p - p) \leq 0, \end{aligned} \quad (2.13)$$

for all $u \in (-1, 1)$ where we used the fact that $p \in (1, 2]$ twice and, without loss of generality, assumed that $C_p \leq p/2$.

The case $p \in (2, \infty)$ can be treated in an analogous manner. Observe that reverse inequalities in (2.12) and (2.13) hold for sufficiently large $C_p > 0$, therefore the auxiliary function F has a non-positive maximum at $u = 1$ and $F(u) \leq 0$ for all $u \in [-1, 1]$.

b) Suppose $p \in (1, 2]$. By [10, Lemma 3.8, Eq. (3.10)], there exists $C_p > 0$ sufficiently large, such that the inequality

$$|x-t|^p - (1-t)^{p-1}(|x|^p - t) \leq C_p t^{p/2} |x-1|^p \quad (2.14)$$

holds for all $t \in [0, 1]$ and $x \in \mathbb{R}$. We shall also make use of the elementary inequality [10, Lemma 3.8, Eq. (3.12)]

$$2^{s-1}(a^s + b^s) \leq (a+b)^s \leq a^s + b^s,$$

which is valid for all $a, b \geq 0$ and $s \in (0, 1]$. Denoting $z = re^{i\theta}$ and $u := \cos \theta$, from the latter inequalities with $s = p/2 \in (0, 1]$, we infer

$$|z-t|^2 = |r-t|^2 + 2rt(1-u) \implies |z-t|^p \leq |r-t|^p + (2rt)^{p/2}(1-u)^{p/2}, \quad (2.15)$$

$$|z-1|^2 = |r-1|^2 + 2r(1-u) \implies |z-1|^p \geq 2^{p/2-1} \left(|r-1|^p + (2r)^{p/2}(1-u)^{p/2} \right). \quad (2.16)$$

Without loss of generality, suppose additionally that $C_p \geq 1$. Then estimates (2.15), (2.14), and (2.16) imply

$$\begin{aligned} |z-t|^p - (1-t)^{p-1}(|z|^p - t) &\leq |r-t|^p - (1-t)^{p-1}(r^p - t) + (2rt)^{p/2}(1-u)^{p/2} \\ &\leq C_p t^{p/2} \left(|r-1|^p + (2r)^{p/2}(1-u)^{p/2} \right) \leq 2^{1-p/2} C_p t^{p/2} |z-1|^p. \end{aligned}$$

If $p \in (2, \infty)$, once more, all the inequalities (2.14), (2.15), and (2.16) are reversed, and the inverted final estimates hold, provided that $C_p > 0$ is chosen sufficiently small. The proof of Lemma 2.13 is complete. $\square$

We are now in a position to prove Theorem 2.3. Suppose $\mathbf{g} \in H^1$ is a non-decreasing sequence such that $\mathbf{g}_n > 0$ for all $n \geq 1$ and $u \in \mathcal{H}_0^1$. We will show that for all $n \geq 2$, we have the inequalities

$$\mathcal{R}(q; \mathbf{g}, u)_{n-1} \lesssim |\nabla u_n|^p - \left(\frac{|u_n|^p}{\mathbf{g}_n^{p-1}} - \frac{|u_{n-1}|^p}{\mathbf{g}_{n-1}^{p-1}} \right) (\nabla \mathbf{g}_n)^{p-1} \lesssim \mathcal{R}(p; \mathbf{g}, u)_{n-1}, \quad (2.17)$$

where the remainders are given by the formula (2.4) and the unspecified constant is independent of n . Next, consider $V_n \geq 0$ for all $n \geq 1$. Multiplying both sides of (2.17) by V_n and summing over n from 2 to ∞ , we arrive at

$$\sum_{n=2}^{\infty} V_n |\nabla u_n|^p + \sum_{n=1}^{\infty} V_{n+1} (\nabla \mathbf{g}_{n+1})^{p-1} \frac{|u_n|^p}{\mathbf{g}_n^{p-1}} - \sum_{n=2}^{\infty} V_n (\nabla \mathbf{g}_n)^{p-1} \frac{|u_n|^p}{\mathbf{g}_n^{p-1}} \left\{ \begin{array}{l} \gtrsim \sum_{n=2}^{\infty} V_n \mathcal{R}(q; \mathbf{g}, u)_{n-1}, \\ \lesssim \sum_{n=2}^{\infty} V_n \mathcal{R}(p; \mathbf{g}, u)_{n-1}. \end{array} \right.$$

Moreover, a shift of the index on the right-hand side reveals

$$\sum_{n=2}^{\infty} V_n |\nabla u_n|^p + \sum_{n=2}^{\infty} \frac{\operatorname{div}(V(\nabla \mathbf{g})^{p-1})_n}{\mathbf{g}_n^{p-1}} |u_n|^p + V_2 (\nabla \mathbf{g}_2)^{p-1} \frac{|u_1|^p}{\mathbf{g}_1^{p-1}} \left\{ \begin{array}{l} \gtrsim \sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(q; \mathbf{g}, u)_n, \\ \lesssim \sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(p; \mathbf{g}, u)_n. \end{array} \right.$$

Recalling $u_0 = \mathbf{g}_0 = 0$ and $\mathbf{g}_1 > 0$, we infer that

$$V_1 |\nabla u_1|^p = V_1 (\nabla \mathbf{g}_1)^{p-1} \frac{|u_1|^p}{\mathbf{g}_1^{p-1}}.$$

Adding and subtracting this term to the left-hand side of the latter inequality yields Theorem 2.3.

To complete the proof, it remains to verify the inequalities stated in (2.17). We shall confine our attention to the case $p \in (1, 2]$. The case $p \in (2, \infty)$ can be handled in a completely analogous fashion; the only modification required is the reversal of the inequalities in (2.10), (2.18), and (2.19). For $p \in (1, 2]$, inequalities (2.17) read

$$|\nabla u_n|^p - \left(\frac{|u_n|^p}{\mathbf{g}_n^{p-1}} - \frac{|u_{n-1}|^p}{\mathbf{g}_{n-1}^{p-1}} \right) (\nabla \mathbf{g}_n)^{p-1} \begin{cases} \gtrsim \left| \sqrt{\frac{\mathbf{g}_{n-1}}{\mathbf{g}_n}} u_n - \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n-1}}} u_{n-1} \right|^2 \left(|\nabla u_n| + \frac{|u_{n-1}|}{\mathbf{g}_{n-1}} |\nabla \mathbf{g}_n| \right)^{p-2}, \\ \lesssim \left| \sqrt{\frac{\mathbf{g}_{n-1}}{\mathbf{g}_n}} u_n - \sqrt{\frac{\mathbf{g}_n}{\mathbf{g}_{n-1}}} u_{n-1} \right|^p. \end{cases} \quad (2.18)$$

We first consider the case $u_{n-1} = 0$. If in addition $u_n = 0$, then (2.18) holds trivially. If, on the other hand, $u_n \neq 0$, then (2.18) stems from the chain of inequalities

$$(p-1)t \leq 1 - (1-t)^{p-1} \leq t \leq t^{p/2}, \quad (2.19)$$

which are valid for every $t \in [0, 1]$, upon substituting $t = \mathbf{g}_{n-1}/\mathbf{g}_n \in (0, 1]$, where the latter inclusion is ensured by the monotonicity of $\mathbf{g}$ (note that the first inequality in (2.19) is an application of the Bernoulli inequality).

Next, assume $u_{n-1} \neq 0$ and temporarily set $\varphi_n := u_n/\mathbf{g}_n$ and substitute for $z = \varphi_n/\varphi_{n-1} \in \mathbb{C}$ and $t = \mathbf{g}_{n-1}/\mathbf{g}_n \in (0, 1]$ into (2.10). It yields the inequalities

$$\left| \frac{\varphi_n}{\varphi_{n-1}} - \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right|^p - \left(1 - \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right)^{p-1} \left(\left| \frac{\varphi_n}{\varphi_{n-1}} \right|^p - \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right) \begin{cases} \gtrsim \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \left| \frac{\varphi_n}{\varphi_{n-1}} - 1 \right|^2 \left(\left| \frac{\varphi_n}{\varphi_{n-1}} - \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right| + 1 - \frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right)^{p-2}, \\ \lesssim \left(\frac{\mathbf{g}_{n-1}}{\mathbf{g}_n} \right)^{p/2} \left| \frac{\varphi_n}{\varphi_{n-1}} - 1 \right|^p \end{cases}$$

Finally, multiplying the latter inequalities by $\mathbf{g}_n^p |\varphi_{n-1}|^p$ implies (2.18). The proof of Theorem 2.3 is complete. $\square$

2.1.2. Proof of Theorem 2.7. The proof proceeds by induction on $\ell \in \mathbb{N}$. The base case $\ell = 1$ is precisely Theorem 2.3 in the particular case $V \equiv 1$. Indeed, in view of definition (1.5) of the discrete gradient, assumption (A1) implies that $\mathbf{g} \in H^1$ defines a non-decreasing sequence; in fact, it is even increasing on $\{\ell-1, \ell, \dots\}$.

Suppose $\ell \geq 2$. Recalling definitions of discrete gradient and divergence (1.5) and our convention, where $\operatorname{div} \nabla = \nabla \operatorname{div}$, specified in Section 1, we have

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p = \sum_{n=\ell-1}^{\infty} |\nabla^{\ell-1} \operatorname{div} u_n|^p \quad \text{for any } u \in \mathcal{H}_0^\ell.$$

Given $\mathbf{g} \in H^\ell$, such that assumptions (A1) and (A2) are fulfilled, we infer that $\operatorname{div} \mathbf{g} = S \nabla \mathbf{g} \in H^{\ell-1}$, where S is the forward shift operator (1.6), satisfies these assumptions for $\ell-1$. The induction hypothesis, applied to $\operatorname{div} u \in \mathcal{H}_0^{\ell-1}$, with $\operatorname{div} \mathbf{g}$, yields

$$\sum_{k=1}^{\ell-1} R_k^{(\ell-1)}(q; \operatorname{div} \mathbf{g}, \operatorname{div} u) \lesssim \sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p - \sum_{n=\ell-1}^{\infty} \frac{-\Delta_p^{(\ell-1)} \operatorname{div} \mathbf{g}_n}{(\operatorname{div} \mathbf{g}_n)^{p-1}} |\operatorname{div} u_n|^p \lesssim \sum_{k=1}^{\ell-1} R_k^{(\ell-1)}(p; \operatorname{div} \mathbf{g}, \operatorname{div} u), \quad (2.20)$$

where the remainders are defined by formulas (2.7) and (2.4). Evidently, for all $k \in \{1, \dots, \ell-1\}$ and $s > 1$, we have the identity

$$R_k^{(\ell-1)}(s; \operatorname{div} \mathbf{g}, \operatorname{div} u) = R_{k+1}^{(\ell)}(s; \mathbf{g}, u), \quad \text{thus} \quad \sum_{k=1}^{\ell-1} R_k^{(\ell-1)}(s; \operatorname{div} \mathbf{g}, \operatorname{div} u) = \sum_{k=2}^{\ell} R_k^{(\ell)}(s; \mathbf{g}, u).$$

Next, we shift the index n in the second sum of the middle part of the latter inequality, arriving at

$$\sum_{n=\ell-1}^{\infty} \frac{-\Delta_p^{(\ell-1)} \operatorname{div} \mathbf{g}_n}{(\operatorname{div} \mathbf{g}_n)^{p-1}} |\operatorname{div} u_n|^p = \sum_{n=1}^{\infty} \frac{S^{\ell-1}(-\Delta_p^{(\ell-1)} \nabla \mathbf{g})_n}{(S^{\ell-1} \nabla \mathbf{g}_n)^{p-1}} |\nabla S^{\ell-1} u_n|^p$$

and we apply Theorem 2.3 with $S^{\ell-1} u \in \mathcal{H}_0^1$, non-decreasing sequence $S^{\ell-1} \mathbf{g} \in H^1$, and the weight

$$V_n := \frac{S^{\ell-1}(-\Delta_p^{(\ell-1)} \nabla \mathbf{g})_n}{(S^{\ell-1} \nabla \mathbf{g}_n)^{p-1}},$$

which is non-negative for all $n \geq 1$ by assumption (A2), obtaining

$$\sum_{n=1}^{\infty} V_n |\nabla S^{\ell-1} u_n|^p + \sum_{n=1}^{\infty} \frac{\operatorname{div}(V(\nabla S^{\ell-1} \mathbf{g})^{p-1})_n}{(S^{\ell-1} \mathbf{g}_n)^{p-1}} |S^{\ell-1} u_n|^p \begin{cases} \gtrsim \sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(q; S^{\ell-1} \mathbf{g}, S^{\ell-1} u)_n, \\ \lesssim \sum_{n=1}^{\infty} V_{n+1} \mathcal{R}(p; S^{\ell-1} \mathbf{g}, S^{\ell-1} u)_n. \end{cases}$$

Let us examine the prefactor in the second sum in greater detail. We observe that

$$\operatorname{div}(V(\nabla S^{\ell-1} \mathbf{g})^{p-1}) = \operatorname{div} \left(\frac{S^{\ell-1}(-\Delta_p^{(\ell-1)} \nabla \mathbf{g})}{(S^{\ell-1} \nabla \mathbf{g})^{p-1}} (S^{\ell-1} \nabla \mathbf{g})^{p-1} \right) = -S^{\ell-1}(-\Delta_p^{(\ell)} \mathbf{g}).$$

Recalling also the definition of the remainders (2.7), the previous inequality rewrites as

$$R_1^{(\ell)}(q; \mathbf{g}, u) \lesssim \sum_{n=\ell-1}^{\infty} \frac{-\Delta_p^{(\ell-1)} \operatorname{div} \mathbf{g}_n}{(\operatorname{div} \mathbf{g}_n)^{p-1}} |\operatorname{div} u_n|^p - \sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell)} \mathbf{g}_n}{\mathbf{g}_n^{p-1}} |u_n|^p \lesssim R_1^{(\ell)}(p; \mathbf{g}, u). \quad (2.21)$$

Note that all the unspecified constants that define the estimates are the same and are determined by the fundamental inequalities (2.10), resp. (2.11).

Altogether, inequalities (2.20) and (2.21) combine to the desired estimate

$$\sum_{k=1}^{\ell} R_k^{(\ell)}(q; \mathbf{g}, u) \lesssim \sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p - \sum_{n=\ell}^{\infty} \frac{-\Delta_p^{(\ell)} \mathbf{g}_n}{\mathbf{g}_n^{p-1}} |u_n|^p \lesssim \sum_{k=1}^{\ell} R_k^{(\ell)}(p; \mathbf{g}, u),$$

which concludes the proof of Theorem 2.7. $\square$

2.1.3. Proof of Theorem 2.9. Recalling the definition of the remainders in (2.7), assumption (A2) guarantees their non-negativity. Furthermore, assumption (A3), i.e. the non-negativity of the sequence

$$\rho_n(\mathbf{g}) := -\Delta_p^{(\ell)} \mathbf{g}_n / \mathbf{g}_n^{p-1} \quad \text{for all } n \geq \ell,$$

combined with the lower bound provided by the abstract higher-order p -Hardy inequality (2.6) established in Theorem 2.7, implies that $\rho(\mathbf{g})$ is indeed a p -Hardy weight of order ℓ .

In order to prove the non-attainability in $\mathcal{H}_0^\ell$, suppose that $u \in \mathcal{H}_0^\ell$ fulfills the p -Birman inequality (2.1) as equality, that is

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p = \sum_{n=\ell}^{\infty} \rho_n(\mathbf{g}) |u_n|^p.$$

Since assumptions (A1) and (A2) are satisfied, Theorem 2.7 implies that all the remainder terms on the left-hand side of inequality (2.6) vanish, i.e. $R_k^{(\ell)}(q; \mathbf{g}, u) = 0$ for all $k \in \{1, \dots, \ell\}$. In particular, for $k = \ell$, we derive a necessary condition (which is, in fact, sufficient)

$$\mathcal{R}(q; \nabla^{\ell-1} \mathbf{g}, \nabla^{\ell-1} u)_n = 0 \quad \text{for all } n \geq \ell;$$

see definition (2.7). If $p \in (1, 2)$ (and hence $q \in (2, \infty)$), assumption (A1) implies in addition that both $\nabla^{\ell-1} \mathbf{g}_n$ and $\nabla^\ell \mathbf{g}_n$ are strictly positive for all $n \geq \ell$. Thus, the vanishing of the second factor in formula (2.4) yields $\nabla^{\ell-1} u_n = \nabla^{\ell-1} u_{n+1} = 0$. In any case, we conclude that the sequence $u \in \mathcal{H}_0^\ell$ satisfies the difference equation of order ℓ

$$\sqrt{\frac{\nabla^{\ell-1} \mathbf{g}_n}{\nabla^{\ell-1} \mathbf{g}_{n+1}}} \nabla^{\ell-1} u_{n+1} - \sqrt{\frac{\nabla^{\ell-1} \mathbf{g}_{n+1}}{\nabla^{\ell-1} \mathbf{g}_n}} \nabla^{\ell-1} u_n = 0 \quad \text{for all } n \geq \ell,$$

with the boundary conditions $u_1 = \dots = u_{\ell-1} = 0$.

However, this equation has the only solution $\mathbf{g}$ up to a multiplicative constant, i.e. $u_n = c \mathbf{g}_n$ for some constant $c \in \mathbb{C}$ and all $n \geq \ell$. By assumption (A1) for $k = 0$, $\mathbf{g}_n > 0$ for all $n \geq \ell$; therefore, the sequence u has compact support if and only if $c = 0$, or equivalently $u \equiv 0$. The proof of Theorem 2.9 is complete. $\square$

2.1.4. Proof of Theorem 2.10. The proof is organized into three distinct parts, in which we demonstrate that the weight $\rho(\mathbf{g}) = -\Delta_p^{(\ell)} \mathbf{g} / \mathbf{g}^{p-1}$ satisfies the properties required for a sequence to be optimal in the sense of Definition 2.1: criticality, non-attainability, and optimality near infinity. The proof is inspired by the linear case $p = 2$ in [32].

The key idea of the proof of non-attainability is that, for any $u \in \mathcal{H}_0^\ell$, equality in the p -Birman inequality entails that all the remainder terms appearing on the left-hand side of inequality (2.6) must vanish. On the other hand, the verification of criticality and optimality near infinity is obtained by substituting a suitable regularization of the parameter sequence $\mathbf{g}$ into inequality (2.6). Specifically, we apply inequality (2.6) to sequences $\{u^N\}_{N \geq 2} \subset \mathcal{H}_0^\ell$ and analyze the asymptotic behavior, as $N \rightarrow \infty$, of the remainder terms $R_k^{(\ell)}(p; \mathbf{g}, u^N)$ appearing on the right-hand side. To this end, we recall the following lemmas from [32]; the proofs can be found therein.

Lemma 2.14. *Let $k \in \mathbb{N}$, $\lambda \in \mathbb{R} \setminus \mathbb{N}_0$, and g be a sequence admitting the asymptotic expansion*

$$g_n = \sum_{j=0}^k a_j n^{\lambda-j} + \mathcal{O}(n^{\lambda-k-1}), \quad \text{as } n \rightarrow \infty, \quad (2.22)$$

for some $a_j \in \mathbb{R}$ and $a_0 \neq 0$. Then the following claims hold true:

(i) For all $m \in \mathbb{Z}$, there exist coefficients $a_j^{(m)} \in \mathbb{R}$ such that

$$S^m g_n = a_0 n^\lambda + \sum_{j=1}^k a_j^{(m)} n^{\lambda-j} + \mathcal{O}(n^{\lambda-k-1}), \quad \text{as } n \rightarrow \infty.$$

(ii) For all $m \in \mathbb{N}_0$, there exist coefficients $b_j^{(m)} \in \mathbb{R}$ with $b_0^{(m)} \neq 0$ such that

$$M^m g_n = \sum_{j=0}^k b_j^{(m)} n^{\lambda-j} + \mathcal{O}(n^{\lambda-k-1}), \quad \text{as } n \rightarrow \infty.$$

(iii) For all $m \in \{0, \dots, k\}$, there exist coefficients $c_j^{(m)} \in \mathbb{R}$ with $c_m^{(m)} \neq 0$ such that

$$\operatorname{div}^m g_n = \sum_{j=m}^k c_j^{(m)} n^{\lambda-j} + \mathcal{O}(n^{\lambda-k-1}), \quad \text{as } n \rightarrow \infty.$$

Claim (iii) of Lemma 2.14 asserts that the discrete derivative $\operatorname{div} \mathbf{g}$, for $\mathbf{g} \in C(\mathbb{Z})$ of the form (2.22), can be treated in a manner analogous to the continuous derivative $\mathbf{g}'$, once $\mathbf{g}$ is naturally extended to a function on $(0, \infty)$, as far as the asymptotic behavior is concerned. The following auxiliary claim, which generalizes the classical *Mean Value Theorem*, further elucidates the relationship between discrete and continuous derivatives.

Lemma 2.15. *Let $n \in \mathbb{Z}$, $m \in \mathbb{N}$, and $g \in C([n, n+m])$ such that $g \in C^m((n, n+m))$. Then there exists $\xi \in (n, n+m)$ such that*

$$\operatorname{div}^m g_n = g^{(m)}(\xi),$$

where we denoted $g_n := g(n)$.

Finally, we will make use of the *Leibniz formula* for the discrete divergence (1.5).

Lemma 2.16. *For all $m \in \mathbb{N}_0$ and sequences $f, g \in C(\mathbb{Z})$, we have*

$$\operatorname{div}^m(fg) = \sum_{j=0}^m \binom{m}{j} (\operatorname{div}^j M^{m-j} f) (\operatorname{div}^{m-j} M^j g),$$

where the multiplication of sequences is to be understood point-wise.

a) *Proof of criticality:* Suppose that $\tilde{\rho}$ is a p -Hardy weight of order ℓ such that $\tilde{\rho}_n \geq \rho_n(\mathbf{g})$ for all $n \geq \ell$. Employing the upper estimate in (2.6) for the weight $\rho(\mathbf{g})$ and the discrete p -Birman inequality (2.1) for the weight $\tilde{\rho}$, we find that

$$0 \leq \sum_{n=\ell}^{\infty} (\tilde{\rho}_n - \rho_n(\mathbf{g})) |u_n|^p \lesssim \sum_{k=1}^{\ell} R_k^{(\ell)}(p; \mathbf{g}, u) \quad (2.23)$$

for all sequences $u \in \mathcal{H}_0^\ell$.

Quick observation of the remainders determined by formula (2.7) reveals that they are all simultaneously annihilated if we set $u = \mathbf{g}$, i.e. $R_k^{(\ell)}(p; \mathbf{g}, \mathbf{g}) \equiv 0$ for all $k \in \{1, \dots, \ell\}$ and $p > 1$. Unfortunately, we cannot directly substitute $u = \mathbf{g}$ to (2.23) since $\mathbf{g} \notin \mathcal{H}_0^\ell$. To this end, we introduce the following regularization of $\mathbf{g}$. First, let us define a smooth cut-off function ξ^N , for any $N \geq 2$, as

$$\xi^N(x) := \begin{cases} 1 & \text{if } x \leq N, \\ \eta\left(\frac{2 \log N - \log x}{\log N}\right) & \text{if } N < x \leq N^2, \\ 0 & \text{if } x > N^2, \end{cases} \quad (2.24)$$

where η is a smooth function, such that $\eta \equiv 0$ on $(-\infty, \varepsilon)$ and $\eta \equiv 1$ on $(1 - \varepsilon, \infty)$ for an arbitrary fixed $\varepsilon \in (0, 1/2)$. We set $u^N := \xi^N \mathbf{g}$ for all $N \geq 2$, where $\xi_n^N := \xi^N(n)$ for all $n \geq 1$.

Notice that by definition (2.24), we have $\xi^N \rightarrow 1$, and therefore $u^N \rightarrow \mathbf{g}$, point-wise as $N \rightarrow \infty$. We will show that this regularization admits the estimates

$$R_k^{(\ell)}(p; \mathbf{g}, u^N) \lesssim \frac{1}{\log^{p-1} N} \quad \text{if } p \in (1, 2] \quad \text{and} \quad R_k^{(\ell)}(p; \mathbf{g}, u^N) \lesssim \frac{1}{\log N} \quad \text{if } p \in (2, \infty) \quad (2.25)$$

for all $k \in \{1, \dots, \ell\}$. Here $\lesssim$ means inequality $\leq$ up to a N -independent constant, which may depend on k or p . Fatou's lemma, together with the inequality (2.23), then implies

$$\sum_{n=\ell}^{\infty} (\tilde{\rho}_n - \rho_n(\mathbf{g})) \mathbf{g}_n^p = \sum_{n=\ell}^{\infty} (\tilde{\rho}_n - \rho_n(\mathbf{g})) \lim_{N \rightarrow \infty} |u_n^N|^p \leq \liminf_{N \rightarrow \infty} \sum_{n=\ell}^{\infty} (\tilde{\rho}_n - \rho_n(\mathbf{g})) |u_n^N|^p = 0.$$

Bearing in mind that all the terms in the sum on the left are non-negative and $\mathbf{g}_n > 0$ for all $n \geq \ell$ by (A1), we may conclude that $\tilde{\rho}_n = \rho_n(\mathbf{g})$ for all $n \geq \ell$, which proves the criticality of the p -Birman weight $\rho(\mathbf{g})$.

It remains to establish the estimates (2.25), which will be carried out in the remainder of part a). Recall that the remainder terms are defined by formula (2.7). Upon replacing the gradient by the discrete divergence, we obtain

$$R_{k+1}^{(\ell)}(p; \mathbf{g}, u^N) = \sum_{n=\ell-k}^{\infty} \frac{-\Delta_p^{(\ell-k-1)} \operatorname{div}^{k+1} \mathbf{g}_n}{(\operatorname{div}^{k+1} \mathbf{g}_n)^{p-1}} \mathcal{R}(p; \operatorname{div}^k \mathbf{g}, \operatorname{div}^k u^N)_n$$

for all $k \in \{0, \dots, \ell-1\}$, and we proceed to analyze the asymptotic behavior of the summands as $n \rightarrow \infty$, with $N \geq 2$ fixed.

We begin by deducing an upper bound of the first factor

$$\frac{-\Delta_p^{(\ell-k-1)} \operatorname{div}^{k+1} \mathbf{g}_n}{(\operatorname{div}^{k+1} \mathbf{g}_n)^{p-1}} \lesssim n^{p(k-\ell+1)} \quad (2.26)$$

for any $n \geq \ell-k$ and $k \in \{0, \dots, \ell-1\}$, where the unspecified constant is n -independent. By assumption (A4) and claim (iii) of Lemma 2.14, we observe that

$$\operatorname{div}^m \mathbf{g}_n = c_0 n^{\ell-m-1/p} \left(1 + \sum_{j=m+1}^{2\ell} \frac{c_j}{c_0} n^{m-j} + \mathcal{O}(n^{m-2\ell-1}) \right), \quad \text{as } n \rightarrow \infty,$$

where $c_0 \neq 0$, for any $m \in \{0, \dots, 2\ell\}$. Furthermore, the generalized binomial theorem implies

$$(\operatorname{div}^m \mathbf{g}_n)^{p-1} = \sum_{j=0}^{2\ell-m} \tilde{c}_j n^{(p-1)(\ell-m-1/p)-j} + \mathcal{O}(n^{(p-1)(\ell-m-1/p)-2\ell+m-1}), \quad \text{as } n \rightarrow \infty, \quad (2.27)$$

where $\tilde{c}_0 \neq 0$. Using the shift operator (1.6) and the definition of the p -Laplacian (1.7), we may rewrite the numerator as

$$-\Delta_p^{(\ell-k-1)} \operatorname{div}^{k+1} \mathbf{g}_n = (-1)^{\ell+k+1} S^{k-\ell+1} \operatorname{div}^{\ell-k-1} (\operatorname{div}^{\ell} \mathbf{g}_n)^{p-1},$$

hence applying claims (i) and (iii) of Lemma 2.14 to the sequence $(\operatorname{div}^{\ell} \mathbf{g})^{p-1}$, we get

$$-\Delta_p^{(\ell-k-1)} \operatorname{div}^{k+1} \mathbf{g}_n \lesssim n^{k-\ell+1/p} \quad \text{and} \quad (\operatorname{div}^{k+1} \mathbf{g}_n)^{p-1} \approx n^{(p-1)(\ell-k-1/p-1)}.$$

Inequality (2.26) follows immediately.

Next, we examine the asymptotic behavior of the second factor. Namely, we prove that

$$\mathcal{R}(p; \operatorname{div}^k \mathbf{g}, \operatorname{div}^k u^N)_n \lesssim \frac{n^{p(\ell-k-1/p-1)}}{\log^{\min\{p, 2\}} N} \quad (2.28)$$

for all $p > 1$, where the unspecified constant does not depend on n nor N . First, we establish the inequality

$$\left| \sqrt{\frac{\operatorname{div}^k \mathbf{g}_n}{\operatorname{div}^k \mathbf{g}_{n+1}}} \operatorname{div}^k u_{n+1}^N - \sqrt{\frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n}} \operatorname{div}^k u_n^N \right| \lesssim \frac{n^{\ell-k-1/p-1}}{\log N} \quad (2.29)$$

for all $N \geq 2$, $n \geq \ell$, and $k \in \{0, \dots, \ell-1\}$. In fact, this result implicitly appeared in [32, Section 2.4] but for the sake of completeness, we will include the derivation of the inequality nonetheless. By assumption (A4) and claims (i) and (iii) of Lemma 2.14, we find that

$$\sqrt{\frac{\operatorname{div}^k \mathbf{g}_n}{\operatorname{div}^k \mathbf{g}_{n+1}}} \lesssim 1,$$

and therefore the left-hand side of (2.29) can be estimated as

$$\sqrt{\frac{\operatorname{div}^k \mathbf{g}_n}{\operatorname{div}^k \mathbf{g}_{n+1}}} \left| \operatorname{div}^k (\xi^N \mathbf{g})_{n+1} - \frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n} \operatorname{div}^k (\xi^N \mathbf{g})_n \right| \lesssim \left| \operatorname{div}^k (\xi^N \mathbf{g})_{n+1} - \frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n} \operatorname{div}^k (\xi^N \mathbf{g})_n \right|.$$

Invoking the Leibniz formula from Lemma 2.16, the last expression rewrites as

$$\begin{aligned} & \left| \sum_{j=0}^k \binom{k}{j} (\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g})_{n+1} (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_{n+1} - \frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n} \sum_{j=0}^k \binom{k}{j} (\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g})_n (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_n \right| \\ & \leq \sum_{j=0}^k \binom{k}{j} (\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g})_{n+1} \left| (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_{n+1} - \frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n} \frac{\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g}_n}{\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g}_{n+1}} (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_n \right|. \end{aligned}$$

Yet another application of (A4) and formulas from Lemma 2.14 reveals that

$$(\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g})_{n+1} \lesssim n^{\ell-j-1/p} \quad \text{and} \quad \frac{\operatorname{div}^k \mathbf{g}_{n+1}}{\operatorname{div}^k \mathbf{g}_n} \frac{\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g}_n}{\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g}_{n+1}} = \frac{1 + \mathcal{O}(1/n)}{1 + \mathcal{O}(1/n)} \frac{1 + \mathcal{O}(1/n)}{1 + \mathcal{O}(1/n)} = 1 + p_n^{(k,j)},$$

where $p_n^{(k,j)} = \mathcal{O}(1/n)$, as $n \rightarrow \infty$, i.e. $p_n^{(k,j)} \lesssim 1/n$, with $p^{(k,k)} \equiv 0$. Altogether, we deduce the upper bound of the left-hand side of (2.29) in the form

$$\sum_{j=0}^k \binom{k}{j} n^{\ell-j-1/p} \left(\left| (\operatorname{div}^{k-j+1} \mathbf{M}^j \xi^N)_{n+1} \right| + \left| p_n^{(k,j)} (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_n \right| \right) \quad (2.30)$$

for all $N \geq 2$, $n \geq \ell$, and $k \in \{0, \dots, \ell-1\}$. In order to obtain the desired estimate (2.29), we have to inspect also the asymptotic behavior of the terms with $\operatorname{div}^m \xi$ in (2.30). To this end, note that for all $x \in [N, N^2]$, we have

$$(\xi^N)'(x) = -\eta' \left(\frac{2 \log N - \log x}{\log N} \right) \frac{1}{x \log N},$$

thus $|(\xi^N)'(x)| \lesssim 1/(x \log N)$, where we estimated η' with $\max_{x \in [0,1]} |\eta'(x)|$. It is straightforward to generalize this bound to higher-order derivatives, i.e.

$$|(\xi^N)^{(m)}(x)| \lesssim \frac{1}{x^m \log N}$$

for all $m \in \mathbb{N}$ and $x > 0$. Consequently, Lemma 2.15 implies that

$$|\operatorname{div}^m \xi_n^N| \lesssim \frac{1}{n^m \log N},$$

which is true for any $m, n \in \mathbb{N}$ and $N \geq 2$. By monotonicity of the right-hand side in the variable n , we also have

$$|\operatorname{div}^m \mathbf{M}^j \xi_n^N| \lesssim \frac{1}{n^m \log N} \quad (2.31)$$

for any $j \in \mathbb{N}_0$, from which we infer further upper estimate of (2.30)

$$\sum_{j=0}^{k-1} \binom{k}{j} n^{\ell-j-1/p} \left(\frac{1}{n^{k-j+1} \log N} + \frac{1}{n} \frac{1}{n^{k-j} \log N} \right) + n^{\ell-k-1/p} \frac{1}{n \log N} \lesssim \frac{n^{\ell-k-1/p-1}}{\log N},$$

finally arriving at the (2.29).

If $p \in (1, 2]$, inequality (2.28) follows immediately from (2.29) and definition (2.4) of the corresponding remainders. If $p \in (2, \infty)$, we must also estimate the second factor in the defining formula (2.4). In particular, for all $k \in \{0, \dots, \ell-1\}$, we show that

$$|\operatorname{div}^{k+1} u_n^N| + \frac{|\operatorname{div}^k u_n^N|}{\operatorname{div}^k \mathbf{g}_n} \operatorname{div}^{k+1} \mathbf{g}_n \lesssim n^{\ell-k-1/p-1} \quad \text{for all } n \geq \ell - k. \quad (2.32)$$

For any $k \in \{0, \dots, \ell\}$, Lemma 2.16 implies

$$\operatorname{div}^k u_n^N = \operatorname{div}^k (\xi^N \mathbf{g})_n = \sum_{j=0}^k \binom{k}{j} (\operatorname{div}^j \mathbf{M}^{k-j} \mathbf{g})_n (\operatorname{div}^{k-j} \mathbf{M}^j \xi^N)_n.$$

Using claims (ii) and (iii) of Lemma 2.14, together with the inequality (2.31), we infer that

$$(\operatorname{div}^j M^{k-j} \mathbf{g})_n \lesssim n^{\ell-j-1/p} \quad \text{and} \quad (\operatorname{div}^{k-j} M^j \xi^N)_n \lesssim n^{j-k}$$

for all $N \geq 2$, $n \geq \ell - k$, and $j \in \{0, \dots, k\}$. Note that the second estimate for $j = k$ follows from the obvious inequality $\xi^N \leq 1$. Altogether, we arrive at

$$\operatorname{div}^k u_n^N \lesssim n^{\ell-k-1/p}.$$

Moreover, employing claims (i) and (iii) of Lemma 2.14 once again, we find that

$$\frac{\operatorname{div}^{k+1} \mathbf{g}_n}{\operatorname{div}^k \mathbf{g}_n} \lesssim \frac{1}{n}.$$

Combining the above inequalities yields the estimate (2.32). Recalling definition (2.4), inequalities (2.29) and (2.32) amount to the desired inequality (2.28) for $p > 2$.

Finally, bearing inequalities (2.26) and (2.28) in mind, for all N sufficiently large, we obtain

$$\begin{aligned} R_{k+1}^{(\ell)}(p; \mathbf{g}, u^N) &= \sum_{n=N-k}^{N^2} \frac{-\Delta_p^{(\ell-1-k)} \operatorname{div}^{k+1} \mathbf{g}_n}{(\operatorname{div}^{k+1} \mathbf{g}_n)^{p-1}} \mathcal{R}(p; \operatorname{div}^k \mathbf{g}, \operatorname{div}^k u^N)_n \lesssim \frac{1}{\log^{\min\{p,2\}} N} \sum_{n=N-k}^{N^2} \frac{1}{n} \\ &\lesssim \frac{1}{\log^{\min\{p,2\}} N} \int_N^{N^2} \frac{1}{n} \, dn = \frac{1}{\log^{\min\{p,2\}-1} N} \end{aligned}$$

for all $k \in \{0, \dots, \ell - 1\}$, which coincides with the aforementioned inequalities (2.25). The proof of criticality is complete.

b) *Proof of optimality near infinity*: Fix an arbitrary $M \geq \ell$. Recalling the definition of optimality near infinity (iii), (2.2), and inequality (2.6), we prove that the weight $\rho(\mathbf{g})$ enjoys this property by finding a sequence $\{u^N\}_{N \geq M} \subset \mathcal{H}_0^M$ such that

$$\lim_{N \rightarrow \infty} \frac{\sum_{k=1}^{\ell} R_k^{(\ell)}(p; \mathbf{g}, u^N)}{\sum_{n=\ell}^{\infty} \rho_n(\mathbf{g}) |u_n^N|^p} = 0. \quad (2.33)$$

Once again, the idea is to find a suitable regularization of the parameter sequence $\mathbf{g}$ so that it becomes an element of the space $\mathcal{H}_0^M$. In particular, we set $u^N := \xi^N \mathbf{g}$, where ξ is a smooth bump function determined by

$$\xi^N(x) := \begin{cases} 0 & \text{if } x \leq N, \\ \eta\left(\frac{\log x - \log N}{\log N}\right) & \text{if } N < x \leq N^2, \\ 1 & \text{if } N^2 < x \leq 2N^2, \\ \eta\left(\frac{\log(2N^3) - \log x}{\log N}\right) & \text{if } 2N^2 < x \leq 2N^3, \\ 0 & \text{if } x > 2N^3, \end{cases}$$

where the function η exhibits the same properties as in (2.24). Hence, indeed, $u^N \in \mathcal{H}_0^M$ for all $N \geq M$.

It is straightforward to show that inequality (2.31) still holds, i.e.

$$|M^j \xi_n^N| \lesssim 1 \quad \text{and} \quad |\operatorname{div}^m M^j \xi_n^N| \lesssim \frac{1}{n^m \log N}$$

for all $m, n \in \mathbb{N}$, $j \in \mathbb{N}_0$, and $N \geq M$. Therefore, the same arguments as in the proof of criticality a) apply, and we obtain the estimates (2.25) even with u^N defined above. Subsequently

$$R_{k+1}^{(\ell)}(p; \mathbf{g}, u^N) \lesssim \frac{1}{\log^{\min\{p,2\}} N} \left(\sum_{n=N-k}^{N^2} \frac{1}{n} + \sum_{n=2N^2-k}^{2N^3} \frac{1}{n} \right) \lesssim \frac{1}{\log^{\min\{p,2\}-1} N}$$

for all $k \in \{0, \dots, \ell - 1\}$ and N sufficiently large, where the unspecified constant does not depend on N nor n , but may depend on p and k .

In order to finish the proof, we show that with our choice of u^N , the denominator in (2.33) is bounded from below by a positive constant. The above inequality then yields the desired limit. First, we inspect the asymptotic behavior of the weight $\rho_n(\mathbf{g}) = -\Delta_p^{(\ell)} \mathbf{g}_n / \mathbf{g}_n^{p-1}$ as $n \rightarrow \infty$. Recalling the definition (1.7) of the p -Laplacian, the numerator can be equivalently written as

$$-\Delta_p^{(\ell)} \mathbf{g}_n = (-1)^\ell S^{-\ell} \operatorname{div}^\ell (\operatorname{div}^\ell \mathbf{g}_n)^{p-1}.$$

Therefore, claims (i) and (iii) of Lemma 2.14, together with the expansion (2.27), imply

$$\rho_n(\mathfrak{g}) \gtrsim \frac{n^{1/p-\ell-1}}{n^{(p-1)(\ell-1/p)}} = \frac{1}{n^{\ell p}} \quad (2.34)$$

for all $n \geq \ell$. Bearing (A4) in mind, we conclude that

$$\sum_{n=\ell}^{\infty} \rho_n(\mathfrak{g}) |u_n^N|^p \geq \sum_{n=N^2}^{2N^2} \rho_n(\mathfrak{g}) \mathfrak{g}_n^p \gtrsim \sum_{n=N^2}^{2N^2} \frac{n^{\ell p-1}}{n^{\ell p}} \geq \int_{N^2}^{2N^2} \frac{1}{n} dn = \log 2$$

completing the proof of optimality near infinity.

c) *Proof of non-attainability:* Under the additional condition of strict positivity in assumption (A3), we first extend the lower bounds in (2.6) from $\mathcal{H}_0^\ell$, to the sequences in H^ℓ , for which equality is attained in the p -Birman inequality (2.1) and both sides are finite. To this end, we first formulate the following lemma, which asserts the density of the range of the operator ∇^ℓ .

Lemma 2.17. *The range of the operator $\nabla^\ell|_{\mathcal{H}_0^\ell}$ is dense in $\mathcal{H}^{\ell,p}$.*

Proof (of Lemma 2.17). Recalling definition (1.5) of the discrete gradient, it is obvious that $\nabla^\ell(\mathcal{H}_0^\ell) \subset \mathcal{H}_0^\ell \subset \mathcal{H}^{\ell,p}$. By the Hahn-Banach Theorem, a subspace of a Banach space is dense if and only if its annihilator is trivial. Consequently the density of $\nabla^\ell(\mathcal{H}_0^\ell)$ in $\mathcal{H}^{\ell,p}$ will be established once we prove that if $v \in (\mathcal{H}^{\ell,p})^* \simeq \mathcal{H}^{\ell,q}$ satisfies

$$\langle v, \nabla^\ell u \rangle := \sum_{n=\ell}^{\infty} \overline{v_n} \nabla^\ell u_n = 0 \quad \text{for all } u \in \mathcal{H}_0^\ell, \quad (2.35)$$

then $v \equiv 0$. Moreover, up to the sign, the algebraic adjoint of the discrete gradient is divergence, i.e.

$$\langle v, \nabla u \rangle = \langle -\operatorname{div} v, u \rangle \quad \text{for all } v \in \mathcal{H}^{\ell,q} \text{ and } u \in \mathcal{H}^{\ell,p}.$$

Therefore, by taking $u = \delta_n$, that is $u_m = \delta_{nm}$, where δ is the Kronecker delta, for all $n \geq \ell$, in the condition (2.35), we observe that v solves the linear difference equation of order ℓ with constant coefficients

$$(-1)^\ell \operatorname{div}^\ell v_n = \sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^j v_{n+j} = 0 \quad \text{for all } n \geq \ell.$$

It is straightforward to verify that the fundamental system of solutions for the above linear difference equation is given by the sequences $1, n, \dots, n^{\ell-1}$. Thus, we obtain

$$v_n = \sum_{j=0}^{\ell-1} c_j n^j \quad \text{for all } n \geq \ell$$

and some coefficients $c_j \in \mathbb{C}$. Furthermore, as $v \in \ell^q(\mathbb{Z})$, it follows that $c_j = 0$ for all $j \in \{0, \dots, \ell-1\}$. Combining this with the assumption $v \in H^\ell$, we conclude that $v \equiv 0$. The proof of Lemma 2.17 is complete. $\square$

Now suppose that $u \in H^\ell$ fulfills the p -Birman inequality (2.1) as equality whose (both) sides are finite, i.e.

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p = \sum_{n=\ell}^{\infty} \rho_n(\mathfrak{g}) |u_n|^p < \infty, \quad (2.36)$$

hence, specifically $\nabla^\ell u \in \mathcal{H}^{\ell,p}$. By Lemma 2.17, there exists a sequence $\{u^N\}_{N=1}^{\infty} \subset \mathcal{H}_0^\ell$ such that $\nabla^\ell u^N \rightarrow \nabla^\ell u$ as $N \rightarrow \infty$ in the norm of $\ell^p(\mathbb{Z})$. In particular, the sequence is Cauchy in $\ell^p(\mathbb{Z})$.

Since assumptions (A1) and (A2) are satisfied, the Hardy inequality of order ℓ , valid on $\mathcal{H}_0^\ell$, with the weight $\rho(\mathfrak{g})$, which is non-negative by assumption (A3), implies

$$0 \leq \left\| \sqrt[p]{\rho(\mathfrak{g})} (u^N - u^M) \right\|_p \leq \left\| \nabla^\ell (u^N - u^M) \right\|_p \xrightarrow{N,M \rightarrow \infty} 0.$$

Consequently, the sequence $\{\sqrt[p]{\rho(\mathfrak{g})} u^N\}_{N=1}^{\infty}$ is Cauchy and therefore convergent in $\ell^p(\mathbb{Z})$. Denote by v the ℓ^p -limit of $\sqrt[p]{\rho(\mathfrak{g})} u^N$ as $N \rightarrow \infty$. It is immediate that $v \in \mathcal{H}^{\ell,p}$ and that, for every $n \in \mathbb{Z}$, we have $\sqrt[p]{\rho_n(\mathfrak{g})} u_n^N \rightarrow v_n$ as $N \rightarrow \infty$. Since $\rho_n(\mathfrak{g}) > 0$ for all $n \geq \ell$ by the strict inequalities in assumption (A3), there exists $\tilde{u} \in H^\ell$ such that $v = \sqrt[p]{\rho(\mathfrak{g})} \tilde{u}$ and, in addition, $u_n^N \rightarrow \tilde{u}_n$ as $N \rightarrow \infty$ for all $n \in \mathbb{Z}$.

Moreover, the limits of all the above-mentioned ℓ^p -convergent sequences must coincide with their point-wise limits, i.e.

$$\nabla^\ell u^N \rightarrow \nabla^\ell \tilde{u} \quad \text{and} \quad \sqrt[p]{\rho(\mathfrak{g})} u^N \rightarrow \sqrt[p]{\rho(\mathfrak{g})} \tilde{u}$$

in $\ell^p(\mathbb{Z})$ as $N \rightarrow \infty$. In particular, we have $\nabla^\ell u = \nabla^\ell \tilde{u}$. This, however, implies $u = \tilde{u}$, because the linear operator ∇^ℓ is injective on H^ℓ . Indeed, if w solves the difference equation

$$\nabla^\ell w_n = 0 \quad \text{for all } n \geq \ell,$$

with the boundary condition $w_n = 0$ for all $n < \ell$, corresponding to the assumption $w \in H^\ell$, we recursively obtain $w_n = 0$ for all $n \in \mathbb{Z}$.

Therefore, recalling (2.36), tending $N \rightarrow \infty$ in the lower estimate in (2.6) for u^N , i.e. in

$$0 \leq \sum_{k=1}^{\ell} R_k^{(\ell)}(q; \mathbf{g}, u^N) \lesssim \sum_{n=\ell}^{\infty} |\nabla^\ell u_n^N|^p - \sum_{n=\ell}^{\infty} \rho_n(\mathbf{g}) |u_n^N|^p,$$

yields that the remainders on the left-hand side must vanish as $N \rightarrow \infty$. In view of definition (2.7), for $k = \ell$, Fatou's lemma entails

$$\lim_{N \rightarrow \infty} \mathcal{R}(q; \nabla^{\ell-1} \mathbf{g}, \nabla^{\ell-1} u^N)_n = \mathcal{R}(q; \nabla^{\ell-1} \mathbf{g}, \nabla^{\ell-1} u)_n = 0$$

for all $n \geq \ell$. Note that if $p \in (1, 2)$ (and hence $q \in (2, \infty)$), assumption (A1) implies in addition that both $\nabla^{\ell-1} \mathbf{g}_n$ and $\nabla^\ell \mathbf{g}_n$ are strictly positive for all $n \geq \ell$, hence the first identity remains valid even for $\nabla^{\ell-1} u_n = \nabla^{\ell-1} u_{n+1} = 0$ by inequality (2.18).

Similarly as in the proof of Theorem 2.9 in Section 2.1.3, we find that $u_n = c \mathbf{g}_n$ for some constant $c \in \mathbb{C}$ and all $n \geq \ell$. From the estimate (2.34) and the asymptotic behavior of the parameter sequence $\mathbf{g}$, given by assumption (A4), we obtain

$$\rho_n(\mathbf{g}) \mathbf{g}_n^p \gtrsim \frac{1}{n^{\ell p}} n^{\ell p-1} = \frac{1}{n} \quad \text{for all } n \geq \ell.$$

Finally, we conclude that

$$\sum_{n=\ell}^{\infty} \rho_n(\mathbf{g}) |u_n|^p \gtrsim \sum_{n=\ell}^{\infty} \frac{1}{n} = \infty,$$

whenever $c \neq 0$, thus, necessarily, $u \equiv 0$. The proof of non-attainability and therefore of Theorem 2.10 is complete. $\square$

3. EXPLICIT OPTIMAL p -BIRMAN WEIGHTS

In this section, we implement the abstract framework introduced in Section 2 for the parameter sequences defined by formulas (2.8) and (2.9). Specifically, in view of Theorem 2.10, we focus on verifying assumptions (A1)–(A4), which will allow us to identify the weights

$$\rho_n^{(\ell,p)} := \frac{-\Delta_p^{(\ell)} \mathbf{g}_n^{(\ell,p)}}{(\mathbf{g}_n^{(\ell,p)})^{p-1}} \quad \text{and} \quad \tilde{\rho}_n^{(\ell,p)} := \frac{-\Delta_p^{(\ell)} \tilde{\mathbf{g}}_n^{(\ell,p)}}{(\tilde{\mathbf{g}}_n^{(\ell,p)})^{p-1}}, \quad (3.1)$$

defined for all $n \geq \ell$, as optimal p -Hardy weights of order ℓ .

The verification of assumptions (A1) and (A4) is straightforward. Assumption (A1) follows directly from the generalized Mean Value Theorem 2.15, while assumption (A4) is established using the known asymptotic expansions of the respective parameter sequences.

Conversely, the validation of assumptions (A2) and (A3) is substantially more involved, primarily due to the nonlinear nature of the p -Laplace operator (1.7). Under the assumption that (A1) holds for $k = \ell$, the joint validity of assumptions (A2) and (A3) is equivalently expressed as

$$(-1)^k \operatorname{div}^k (\operatorname{div}^\ell \mathbf{g}_n)^{p-1} \geq 0 \quad \text{for all } n > 0 \text{ and } k \in \{0, \dots, \ell\}, \quad (3.2)$$

which naturally leads to the notion of *completely monotone* functions.

Intermezzo: Completely monotone functions. A function $f : (0, \infty) \rightarrow \mathbb{R}$ is called *completely monotone* if and only if it is of the class C^∞ and it satisfies the *Bernstein–Hausdorff–Widder condition*

$$(-1)^k f^{(k)}(x) \geq 0 \quad \text{for all } k \geq 0 \text{ and } x > 0. \quad (3.3)$$

We denote the set of completely monotone functions as

$$\mathcal{CM} := \{f : (0, \infty) \rightarrow \mathbb{R} \mid f \text{ is completely monotone}\}.$$

The notion of a completely monotone functions dates back to 1914 and S. Bernstein, who studied functions on $(-\infty, 0)$ with non-negative derivatives of all orders, i.e. functions satisfying (3.3) after omitting the alternating sign $(-1)^k$. Such functions are often called *absolutely monotone* in the literature. This, however, coincides with complete monotonicity up to a change of sign in the variable. Since then, the properties of completely monotone

functions have been thoroughly studied as they arise frequently in a wide range of mathematical disciplines, such as the moment problem, probabilistic theory, and mathematical analysis. Here we discuss their surprising relevance also in the context of Hardy inequalities.

The often-cited book by D. V. Widder [30] contains a detailed account of the properties of completely monotone functions. We point out Bernstein's theorem [30, Theorem 12a], which characterizes completely monotone functions as the Laplace transform of a positive Borel measure. Namely, a function $f : (0, \infty) \rightarrow \mathbb{R}$ is completely monotone if and only if there exists a measure μ on $[0, \infty)$ such that

$$f(x) = \int_{[0, \infty)} e^{-xt} d\mu(t) \quad \text{for all } x > 0. \quad (3.4)$$

Remark 3.1. The Bernstein Theorem further implies that

$$(-1)^k f^{(k)}(x) = \int_{[0, \infty)} t^k e^{-xt} d\mu(t) \geq 0 \quad \text{for all } k \geq 0 \text{ and } x > 0,$$

which, in particular, shows that $f^{(k)}(x) \neq 0$ for all $k \geq 0$ and all $x > 0$, unless $f \in \mathcal{CM}$ is constant; see also [27, Remark 1.5]

Moreover, the aforementioned *Hausdorff moment problem* was studied therein. Analogously to the continuous case (3.3), a sequence $\{u_n\}_{n=0}^\infty \subset \mathbb{R}$ is called *completely monotone* if and only if

$$(-1)^k \operatorname{div}^k u_n \geq 0 \quad \text{for all } k \geq 0 \text{ and } n \geq 0.$$

We denote the set of completely monotone sequences on $\mathbb{N}_0$ as

$$\mathcal{CM}(\mathbb{N}_0) := \{u \in C(\mathbb{N}_0) \mid u \text{ is completely monotone}\}.$$

Hausdorff's Theorem [30, Theorem 4a] provides a necessary and sufficient condition for a sequence $\{u_n\}_{n \geq 0}$ to be completely monotone, that it admits the integral representation

$$u_n = \int_0^1 t^n d\mu(t) \quad \text{for all } n \geq 0, \quad (3.5)$$

for some finite Borel measure μ on the interval $[0, 1]$.

This formula also explains the connection between completely monotone functions and sequences. If $f \in \mathcal{CM}$ is right-continuous at 0, i.e. the limit of $f(x)$ as $x \rightarrow 0_+$ is finite, then $f_n := f(n)$ for all $n \geq 0$, where $f(0)$ is defined by the corresponding one-sided limit, defines a completely monotone sequence $f \in \mathcal{CM}(\mathbb{N}_0)$ by Lemma 2.15. On the other hand, if $u \in \mathcal{CM}(\mathbb{N}_0)$, one easily observes that

$$u(x) := \int_0^1 t^x d\mu(t),$$

where the measure μ is given by formula (3.5), defines a completely monotone function $u \in \mathcal{CM}$ continuous at 0. Consequently, this reduces many problems concerning completely monotone sequences to the more tractable continuous setting.

We now recall two classical properties of completely monotone functions. For the sake of completeness, we also include their short proofs. The first proposition asserts that the class $\mathcal{CM}$ is closed under multiplication.

Proposition 3.2. *If $f, g \in \mathcal{CM}$, then also $fg \in \mathcal{CM}$. Equivalently $\mathcal{CM} \cdot \mathcal{CM} \subset \mathcal{CM}$.*

Proof. The proof follows immediately from the Leibniz formula

$$(-1)^k (fg)^{(k)} = \sum_{j=0}^k \binom{k}{j} (-1)^j f^{(j)} (-1)^{k-j} g^{(k-j)}$$

for any $k \geq 0$ and definition (3.3). □

In view of the inequalities (3.2), we are particularly interested in determining whether the positive $(p-1)$ -st power of the functions $\operatorname{div}^\ell \mathbf{g} \in \mathcal{CM}$ (see Lemmas 3.5 and 3.14) is itself completely monotone. However, a positive power of a completely monotone function is, in general, not completely monotone. This stability under positive powers is instead enjoyed by the class of *infinitely divisible* functions. A smooth function $f : (0, \infty) \rightarrow \mathbb{R}$ is called *infinitely divisible* if and only if

$$f^\lambda \in \mathcal{CM} \quad \text{for all } \lambda > 0.$$

We denote the set of infinitely divisible functions by $\mathcal{ID}$.

Infinitely divisible functions are closely related to another class, the *logarithmically completely monotone* functions. A smooth function $f : (0, \infty) \rightarrow (0, \infty)$ is said to be *logarithmically completely monotone* if and only if

$$-(\log f)' \in \mathcal{CM}. \quad (3.6)$$

We denote the set of logarithmically completely monotone functions by $\mathcal{LCM}$.

Proposition 3.3. *A logarithmically completely monotone function is infinitely divisible, and hence completely monotone. Equivalently $\mathcal{LCM} \subset \mathcal{ID} \subset \mathcal{CM}$.*

Proof. Obviously, if $f : (0, \infty) \rightarrow (0, \infty)$ is logarithmically completely monotone, then so is the function f^λ for any $\lambda > 0$. Consequently, it suffices to establish the inclusion $\mathcal{LCM} \subset \mathcal{CM}$.

Let $f \in \mathcal{LCM}$ and define a function $g : (-\infty, 0) \rightarrow (0, \infty) : x \mapsto f(-x)$. Observe that the function f is completely monotone if and only if g is absolutely monotone on $(b, 0)$ for all $b < 0$, i.e.

$$g^{(k)}(x) \geq 0 \quad \text{for all } k \geq 0 \text{ and } x \in (b, 0).$$

Fix arbitrary $b < 0$. By definition (3.6) and the Bernstein–Hausdorff–Widder condition (3.3), the function $(b, 0) \rightarrow (0, \infty) : x \mapsto (\log g(x))' = -(\log f)'(-x)$ is absolutely monotone on $(b, 0)$. Therefore, $\log g : (b, 0) \rightarrow \mathbb{R}$ is a non-decreasing function with a finite limit at the point b , thus there exist a constant $C \in \mathbb{R}$, such that $C + \log g : (b, 0) \rightarrow (0, \infty)$ is absolutely monotone on $(b, 0)$. Finally, we deduce that

$$g(x) = e^{-C} e^{C + \log g(x)}$$

is absolutely monotone on $(b, 0)$ as a constant multiple of a composition of absolutely monotone functions $x \mapsto e^x$ and $x \mapsto C + \log g(x)$. The proof of Proposition 3.3 is complete. $\square$

Remark 3.4. In fact, an even stronger connection between infinitely divisible and logarithmically completely monotone functions was established, for instance, in [27, Theorem 5.9]: a function $f : (0, \infty) \rightarrow (0, \infty)$ is infinitely divisible if and only if it is logarithmically completely monotone.

3.1. Parameter sequence as a ratio of Gamma functions. Recalling definition (2.8), we study the properties of the sequence

$$\mathbf{g}_n^{(\ell, p)} = \frac{\Gamma(n + 1/q)}{\Gamma(n - \ell + 1)} \quad \text{for all } n \in \mathbb{Z}. \quad (3.7)$$

In particular, we verify that the assumptions (A1)–(A4) are satisfied. Consequently, by Theorem 2.10, the sequence $\rho^{(\ell, p)}$, defined by formula (3.1), yields an optimal p -Hardy inequality of order ℓ .

Firstly, since the reciprocal Gamma function $1/\Gamma$ is an entire function on the complex plane with zeros at the points $0, -1, \dots$, it follows that $\mathbf{g}^{(\ell, p)} \in H^\ell$.

We now analyze the behavior of the sequence $\mathbf{g}^{(\ell, p)}$ under the action of the discrete divergence operator defined in (1.5). In particular, we uncover a full analog of the classical differentiation formula for monomials

$$\frac{d^k x^{\ell-1/p}}{dx^k} = (\ell - k + 1/q)_k x^{\ell-k-1/p}.$$

By extending the definition in (3.7) to all $\ell \in \mathbb{Z}$, we obtain the following result.

Lemma 3.5. *Let $\ell \in \mathbb{Z}$ and $k \in \mathbb{N}_0$. Then, for all $n \in \mathbb{Z}$, we have the identity*

$$\operatorname{div}^k \mathbf{g}_n^{(\ell, p)} = (\ell - k + 1/q)_k \mathbf{g}_n^{(\ell-k, p)}. \quad (3.8)$$

Proof. Bearing the operator identity $\operatorname{div} = S - I$ in mind, we obtain

$$\operatorname{div}^k \mathbf{g}_n^{(\ell, p)} = (-1)^k \sum_{j=0}^k \binom{k}{j} (-1)^j \frac{\Gamma(n + j + 1/q)}{\Gamma(n - \ell + j + 1)}$$

for all $n \in \mathbb{Z}$. Using the recurrence relation $\Gamma(z + 1) = z\Gamma(z)$, valid for all $z \in \mathbb{C}$, we can further rewrite this expression in terms of the *regularized hypergeometric function* (see [9, section 15]) as

$$\operatorname{div}^k \mathbf{g}_n^{(\ell, p)} = (-1)^k \Gamma(n + 1/q) \sum_{j=0}^k \frac{(-k)_j (n + 1/q)_j}{\Gamma(n - \ell + 1 + j) j!} = (-1)^k \Gamma(n + 1/q) {}_2F_1 \left(\begin{matrix} -k, n + 1/q \\ n - \ell + 1 \end{matrix} \middle| 1 \right).$$

Moreover, since $k \in \mathbb{N}_0$, the *Chu–Vandermonde identity* [9, Eq. (15.4.24)] implies

$${}_2F_1 \left(\begin{matrix} -k, n + 1/q \\ n - \ell + 1 \end{matrix} \middle| 1 \right) = \frac{(-\ell - 1/q + 1)_k}{\Gamma(n - \ell + k + 1)},$$

hence, in view of the definition (3.7), we conclude that

$$\operatorname{div}^k \mathfrak{g}_n^{(\ell,p)} = (-1)^k (-\ell - 1/q + 1)_k \frac{\Gamma(n + 1/q)}{\Gamma(n - \ell + k + 1)} = (\ell - k + 1/q)_k \mathfrak{g}_n^{(\ell-k,p)}$$

for all $n \in \mathbb{Z}$. $\square$

The explicit formula (3.8) plays a crucial role, as it yields a closed-form expression for the sequence $\operatorname{div}^\ell \mathfrak{g}^{(\ell,p)}$, thereby substantially simplifying the subsequent computations. This feature is the main motivation for working with the sequences $\mathfrak{g}^{(\ell,p)}$ rather than with $\tilde{\mathfrak{g}}^{(\ell,p)}$. Lemma 3.5, in turn, directly implies the following corollary.

Corollary 3.6. *Let $\ell \in \mathbb{N}$. Then the sequences $\mathfrak{g}^{(\ell,p)}$, defined by (3.7), satisfy the assumption (A1), i.e.*

$$\nabla^k \mathfrak{g}_n^{(\ell,p)} > 0 \quad \text{for all } n \geq \ell \text{ and } k \in \{0, \dots, \ell\}.$$

Proof. By definitions of the discrete gradient and divergence (1.5), Lemma 3.5 implies

$$\nabla^k \mathfrak{g}_n^{(\ell,p)} = (\ell - k + 1/q)_k \mathfrak{g}_{n-k}^{(\ell-k,p)} > 0$$

for all $n \geq \ell \geq k \geq 0$. The strict positivity follows directly from the defining formula (3.7). $\square$

Moreover, the complete monotonicity of the sequence

$$\operatorname{div}^\ell \mathfrak{g}_n^{(\ell,p)} = \left(\frac{1}{q}\right)_\ell \frac{\Gamma(n + 1/q)}{\Gamma(n + 1)} \quad \text{for all } n \geq 0, \quad (3.9)$$

also follows directly from formula (3.8). In the subsequent proposition, we demonstrate that this sequence, in fact, satisfies even the stronger condition of logarithmic complete monotonicity (3.6). The assertion is a direct consequence of the more general result [1, Theorem 10]. Nevertheless, for the sake of completeness, we include a brief explanatory proof.

Proposition 3.7. *Let $q > 1$. Then the function*

$$G : (0, \infty) \rightarrow (0, \infty) : x \mapsto \frac{\Gamma(x + 1/q)}{\Gamma(x + 1)} \quad (3.10)$$

is logarithmically completely monotone.

Proof. Observe first that $G(x) > 0$ for all $x > 0$. We temporarily introduce the auxiliary function

$$H(x) := -(\log G(x))' = \psi(x + 1) - \psi(x + 1/q) \quad \text{for all } x > 0,$$

where ψ denotes the *digamma function*, which admits the integral representation [9, Eq. (5.9.16)]

$$\psi(x) := \frac{\Gamma'(x)}{\Gamma(x)} = -\gamma + \int_0^\infty \frac{e^{-t} - e^{-xt}}{1 - e^{-t}} dt \quad \text{for all } x > 0, \quad (3.11)$$

and $\gamma \in \mathbb{R}$ denotes the *Euler–Mascheroni constant*. Substituting this representation into the definition of H yields

$$H(x) = \int_0^\infty e^{-xt} \frac{e^{-t/q} - e^{-t}}{1 - e^{-t}} dt,$$

so that H can be written as a Laplace transform of the form (3.4). Since $q > 1$, the density of the associated Borel measure $(e^{-t/q} - e^{-t})/(1 - e^{-t})$ is strictly positive for all $t > 0$. Consequently,

$$(-1)^k H^{(k)}(x) = \int_0^\infty e^{-xt} t^k \frac{e^{-t/q} - e^{-t}}{1 - e^{-t}} dt > 0$$

for all $k \in \mathbb{N}_0$ and all $x > 0$. It follows that $H \in \mathcal{CM}$, and therefore $G \in \mathcal{LCM}$. $\square$

Corollary 3.8. *The sequences $\mathfrak{g}^{(\ell,p)}$, defined by (3.7), satisfy the assumptions (A2) and (A3). In fact, we have the strict inequalities*

$$(-1)^k \operatorname{div}^k (\operatorname{div}^\ell \mathfrak{g}_n^{(\ell,p)})^{p-1} > 0 \quad \text{for all } n \geq 0 \text{ and } k \geq 0. \quad (3.12)$$

Proof. As a direct consequence of Propositions 3.7 and 3.3, the function G , defined in (3.10), belongs to the class $\mathcal{LCM}$, and hence to $\mathcal{ID}$. Recalling also the formula (3.9), the right-continuity of G at zero, together with the Mean Value Theorem 2.15, implies that, for any $\lambda > 0$, we have

$$(-1)^k \operatorname{div}^k (\operatorname{div}^\ell \mathfrak{g}_n^{(\ell,p)})^\lambda > 0 \quad \text{for all } n \geq 0 \text{ and } k \geq 0,$$

where the strict inequality follows from Remark 3.1. In particular, the choice $\lambda = p - 1 > 1$ reproduces inequalities (3.12), which, in turn, guarantee that the assumptions (A2) and (A3) are satisfied. $\square$

Finally, we recall the asymptotic expansion for the ratio of Gamma functions [9, § 5.11(iii)]:

$$\frac{\Gamma(x+a)}{\Gamma(x+b)} \sim x^{a-b} \sum_{k=0}^{\infty} \frac{G_k(a,b)}{x^k}, \quad \text{as } x \rightarrow \infty, \text{ for all } a, b \in \mathbb{R},$$

where the coefficients $G_k(a,b)$ are explicitly determined in terms of *generalized Bernoulli polynomials* $B_k^{(\lambda)}$ [9, § 24.16(i)] as

$$G_k(a,b) := \binom{a-b}{k} B_k^{(a-b+1)}(a) \quad \text{for all } k \in \mathbb{N}_0,$$

in particular, $G_0(a,b) = 1$.

Corollary 3.9. *The sequences $\mathfrak{g}^{(\ell,p)}$, defined by (3.7), satisfy the assumption (A4). In particular, $\mathfrak{g}^{(\ell,p)}$ admits the asymptotic expansion*

$$\mathfrak{g}_n^{(\ell,p)} \sim n^{\ell-1/p} \sum_{k=0}^{\infty} \frac{G_k(1/q, 1-\ell)}{n^k}, \quad \text{as } n \rightarrow \infty.$$

In light of Theorem 2.10, the results established in this section collectively yield the following theorem, which we regard as the second main result of this paper.

Theorem 3.10. *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the sequence*

$$\rho_n^{(\ell,p)} = \frac{-\Delta_p^{(\ell)} \mathfrak{g}_n^{(\ell,p)}}{(\mathfrak{g}_n^{(\ell,p)})^{p-1}} = \left(\frac{p-1}{p} \right)_\ell \frac{1}{(n-\ell+1-1/p)_\ell^{p-1}} \sum_{j=0}^{\ell} \frac{(-\ell)_j (n-\ell+1-1/p)_j^{p-1}}{(n-\ell+1)_j^{p-1}} \frac{1}{j!}, \quad (3.13)$$

for all $n \geq \ell$, where $\mathfrak{g}^{(\ell,p)}$ is given by (3.7), is an optimal discrete p -Hardy weight of order ℓ .

Remark 3.11. If $p \geq 2$ is an integer, the weight $\rho^{(\ell,p)}$ can be represented in terms of *generalized hypergeometric functions* [9, § 16] as

$$\rho_n^{(\ell,p)} = \left(\frac{p-1}{p} \right)_\ell \frac{1}{(n-\ell+1-1/p)_\ell^{p-1}} {}_pF_{p-1} \left(\begin{matrix} -\ell, n-\ell+1-\frac{1}{p}, \dots, n-\ell+1-\frac{1}{p} \\ n-\ell+1, \dots, n-\ell+1 \end{matrix} \middle| 1 \right),$$

for all $n \geq \ell$. In particular, for $p = 2$, an application of Lemma 3.5 yields the further simplification

$$\rho_n^{(\ell,2)} = \left(\frac{1}{2} \right)_\ell \frac{1}{n(n-1/2)(n-1) \cdots (n-\ell+1/2)} > \left(\frac{1}{2} \right)_\ell \frac{1}{n^{2\ell}}, \quad (3.14)$$

for all $n \geq \ell$. Consequently, this provides an additional, more explicit example of an optimal Hardy weight of order ℓ in the linear case $p = 2$, thereby expanding the family of such weights constructed in [32, Section 3].

Remark 3.12. In view of Remark 2.12, we further observe that the parameter sequences

$$\mathfrak{g}_n^{(\ell,p)}(s) := \frac{\Gamma(n+s)}{\Gamma(n-\ell+1)}, \quad \text{for all } n \in \mathbb{Z},$$

generate discrete p -Hardy weights of order ℓ , given by

$$\rho_n^{(\ell,p)}(s) := \frac{-\Delta_p^{(\ell)} \mathfrak{g}_n^{(\ell,p)}(s)}{(\mathfrak{g}_n^{(\ell,p)}(s))^{p-1}} = (s)_\ell^{p-1} \frac{1}{(n-\ell+s)_\ell^{p-1}} \sum_{j=0}^{\ell} \frac{(-\ell)_j (n-\ell+s)_j^{p-1}}{(n-\ell+1)_j^{p-1}} \frac{1}{j!} \quad \text{for all } n \geq \ell,$$

which are critical if $s \in (0, 1/q]$, non-attainable if $s \in [1/q, 1)$, and optimal near infinity if $s = 1/q$.

3.2. Parameter sequence as a linear combination of monomials. Motivated by the specific choice of the parameter sequences in (9), as well as by the optimal first-order p -Hardy inequality (7), we also analyze our second candidate $\tilde{\mathfrak{g}}^{(\ell,p)}$ defined by (2.9).

For $n \in \mathbb{N}$ and $k \in \{0, \dots, n-1\}$, we denote by

$$s(n, k) := (-1)^{n+k} \sum_{1 \leq i_1 < \dots < i_{n-k} < n} i_1 i_2 \dots i_{n-k}$$

the (signed) Stirling numbers of the first kind (see [9, § 26.8]). By convention, we additionally set $s(n, n) := 1$ for all $n \in \mathbb{N}$. Using the combinatorial identity [9, Eq. (26.8.7)], we obtain

$$\tilde{\mathfrak{g}}_n^{(\ell,p)} = n^{1-1/p} \prod_{j=1}^{\ell-1} (n-j) = \sum_{j=1}^{\ell} s(\ell, j) n^{j-1/p} \quad (3.15)$$

for all $n \geq 0$.

It is evident that $\tilde{\mathbf{g}}^{(\ell,p)} \in H^\ell$ automatically satisfies assumption (A4). We now proceed to verify assumption (A1).

Proposition 3.13. *The sequences $\tilde{\mathbf{g}}^{(\ell,p)}$, defined by formula (3.15), satisfy the assumption (A1), i.e.*

$$\nabla^k \tilde{\mathbf{g}}_n^{(\ell,p)} > 0 \quad \text{for all } n \geq \ell \text{ and } k \in \{0, \dots, \ell\}. \quad (3.16)$$

Proof. Notice that since $\tilde{\mathbf{g}}^{(\ell,p)} \in H^\ell$ by (3.15), it suffices to show inequalities (3.16) for $k = \ell$. Indeed, by definition of discrete gradient (1.5), $\nabla^\ell \tilde{\mathbf{g}}_n^{(\ell,p)} > 0$ for all $n \geq \ell$ means that the sequence $\nabla^{\ell-1} \tilde{\mathbf{g}}^{(\ell,p)}$ is strictly increasing on $\{\ell-1, \ell, \dots\}$. Since $\nabla^{\ell-1} \tilde{\mathbf{g}}^{(\ell,p)} \in H^\ell$, hence $\nabla^{\ell-1} \tilde{\mathbf{g}}_{\ell-1}^{(\ell,p)} = 0$, we conclude that $\nabla^{\ell-1} \tilde{\mathbf{g}}_n^{(\ell,p)} > 0$ for all $n \geq \ell$. Iteration of this argument yields (3.16) for all $k \in \{0, \dots, \ell\}$.

In order to finish the proof, we establish the inequality

$$\nabla^\ell \tilde{\mathbf{g}}_n^{(\ell,p)} > 0$$

for all $n \geq \ell$. The key idea is to examine the natural extension of the sequence $\tilde{\mathbf{g}}^{(\ell,p)}$ to all positive real numbers, defined by

$$\tilde{\mathbf{g}}^{(\ell,p)}(x) := x^{1-1/p} \prod_{j=1}^{\ell-1} (x-j), \quad (3.17)$$

and make use of Lemma 2.15, which implies that for any $n \geq \ell$ there exists $\xi \in (n-\ell, n)$, such that

$$\nabla^\ell \tilde{\mathbf{g}}_n^{(\ell,p)} = \frac{d^\ell \tilde{\mathbf{g}}^{(\ell,p)}}{dx^\ell}(\xi).$$

Furthermore, bearing (3.15) in mind, we have

$$\frac{d^\ell \tilde{\mathbf{g}}^{(\ell,p)}}{dx^\ell}(x) = \sum_{j=1}^{\ell} s(\ell, j) \frac{d^\ell}{dx^\ell} x^{j-1/p} = \sum_{j=1}^{\ell} b_j^{(\ell,p)} x^{j-\ell-1/p} \quad (3.18)$$

for any $x > 0$, where

$$b_j^{(\ell,p)} := (-1)^{\ell+j} s(\ell, j) \prod_{k=0}^{\ell-1} |j-k-1/p| > 0$$

for all $j \in \{1, \dots, \ell\}$. The desired inequality follows immediately. $\square$

Recalling the operator formula $\text{div} = S - I$, we define the function

$$\tilde{G} : (0, \infty) \rightarrow \mathbb{R} : x \mapsto \text{div}^\ell \tilde{\mathbf{g}}^{(\ell,p)}(x) := \sum_{j=0}^{\ell} \binom{\ell}{j} (-1)^{\ell+j} \tilde{\mathbf{g}}^{(\ell,p)}(x+j), \quad (3.19)$$

where $\tilde{\mathbf{g}}^{(\ell,p)}$ is given by (3.17). Observing that higher-order derivatives of the function $\tilde{\mathbf{g}}^{(\ell,p)}$ in (3.18) begin to change sign, we arrive at the following lemma.

Lemma 3.14. *The function $\tilde{G} : (0, \infty) \rightarrow (0, \infty)$, defined by formula (3.19), is completely monotone.*

Proof. Consider an arbitrary $k \in \mathbb{N}_0$. Evidently, for any $x > 0$, we have

$$(-1)^k \tilde{G}^{(k)}(x) = (-1)^k \frac{d^k}{dx^k} \text{div}^\ell \tilde{\mathbf{g}}^{(\ell,p)}(x) = (-1)^k \text{div}^\ell \frac{d^k}{dx^k} \tilde{\mathbf{g}}^{(\ell,p)}(x).$$

Therefore, by Lemma 2.15, we readily infer that there exists $\xi \in (x, x+\ell)$ such that

$$(-1)^k \tilde{G}^{(k)}(x) = (-1)^k \frac{d^{\ell+k} \tilde{\mathbf{g}}^{(\ell,p)}}{dx^{\ell+k}}(\xi).$$

Furthermore, recalling formula (3.15), we find that

$$(-1)^k \frac{d^{\ell+k} \tilde{\mathbf{g}}^{(\ell,p)}}{dx^{\ell+k}}(x) = (-1)^k \sum_{j=1}^{\ell} s(\ell, j) \frac{d^{\ell+k}}{dx^{\ell+k}} x^{j-1/p} = \sum_{j=1}^{\ell} c_j^{(\ell,p)} x^{j-\ell-k-1/p}$$

for any $x > 0$, where

$$c_j^{(\ell,p)} = (-1)^k (-1)^{\ell+j+k} s(\ell, j) \prod_{m=0}^{\ell+k-1} |j-m-1/p| > 0$$

for all $j \in \{1, \dots, \ell\}$. Altogether, we conclude that

$$(-1)^k \tilde{G}^{(k)}(x) > 0 \quad \text{for all } k \geq 0 \text{ and } x > 0,$$

hence the function $\tilde{G}$ is completely monotone. $\square$

The generalized Mean Value Theorem 2.15 ensures that $\operatorname{div}^\ell \tilde{\mathbf{g}}^{(\ell,p)} \in \mathcal{CM}(\mathbb{N}_0)$ as well, thus inequalities (3.2) are fulfilled if the exponent $p - 1$ is omitted, in fact

$$(-1)^k \operatorname{div}^k \operatorname{div}^\ell \tilde{\mathbf{g}}_n^{(\ell,p)} > 0 \quad \text{for all } n \geq 0 \text{ and } k \geq 0.$$

In contrast to the function $G \in \mathcal{LCM}$ defined in (3.10), we were unable to establish the logarithmic complete monotonicity of $\tilde{G}$, and consequently we could not derive the inequalities in (3.2) for arbitrary $p > 1$. Nevertheless, by imposing suitable restrictions on the admissible values of $\ell \in \mathbb{N}$ and $p > 1$, we can infer the following result from Lemma 3.14.

Proposition 3.15. *Assume that one of the following conditions is satisfied:*

- (i) $\ell \in \mathbb{N}$ and $p \geq 2$ is an integer,
- (ii) $\ell \in \{3, 4, 5, 6, 7\}$ and $p > 2$,
- (iii) $\ell = 1$ or $\ell = 2$ and $p > 1$.

Then the sequence $\tilde{\mathbf{g}}^{(\ell,p)}$, given by (3.15), satisfies the assumptions (A2) and (A3), i.e.

$$(-1)^k \operatorname{div}^k (\operatorname{div}^\ell \tilde{\mathbf{g}}_n^{(\ell,p)})^{p-1} \geq 0 \quad \text{for all } n \geq 0 \text{ and } k \in \{1, \dots, \ell\}.$$

Proof. The proof is carried out in three separate steps, each corresponding to the fulfillment of the conditions (i), (ii), and (iii), respectively.

a) *Let $\ell \in \mathbb{N}$ and $p \geq 2$ be an integer:* Recall that the set of completely monotone functions is closed under multiplication by Proposition 3.2. Consequently, from Lemma 3.14, we infer that $\tilde{G}^{p-1} \in \mathcal{CM}$ and, by Remark 3.1 and Lemma 2.15, also

$$(-1)^k \operatorname{div}^k (\operatorname{div}^\ell \tilde{\mathbf{g}}_n^{(\ell,p)})^{p-1} > 0 \quad \text{for all } n \geq 0 \text{ and } k \geq 0. \quad (3.20)$$

In particular the inequalities hold for all $k \in \{1, \dots, \ell\}$, which completes the proof of part a).

If $p > 1$ is non-integer, the validation of inequalities (3.2) is much more complicated. In contrast to part a) of this proof, the inequalities will only be established for $k \in \{0, \dots, \ell\}$, for which reason we introduce the following notation; see [28].

Let $N \in \mathbb{N}^* := \mathbb{N} \cup \{\infty\}$, the function $f : (0, \infty) \rightarrow \mathbb{R}$ is called *completely monotone of order N* if and only if it is of the class C^N and

$$(-1)^k f^{(k)}(x) \geq 0 \quad \text{for all } k \in \{0, \dots, N\} \text{ and } x > 0.$$

The set of completely monotone functions of order N is denoted as $\mathcal{CM}^N$, in particular, we have $\mathcal{CM}^\infty = \mathcal{CM}$. Furthermore, we define the set of functions, whose positive power is completely monotone, as

$$\mathcal{CM}_\lambda^N := \{f : (0, \infty) \rightarrow [0, \infty) \mid f^\lambda \in \mathcal{CM}^N\}$$

for any $\lambda > 0$ and $N \in \mathbb{N}^*$. In particular, we have $\mathcal{CM}_1^N = \mathcal{CM}^N$ and we abbreviate the notation if $N = \infty$ as $\mathcal{CM}_\lambda := \mathcal{CM}_\lambda^\infty$.

Invoking Lemma 2.15 once more, we observe that in order to show (3.2), it suffices to establish $\tilde{G} \in \mathcal{CM}_{p-1}^\ell$. To this end, we will make use of the following results from [28, Theorem 4] and [29].

Lemma 3.16. *The following inclusions hold true:*

- (i) For any $N \in \{0, \dots, 5\}$, we have $\mathcal{CM}_\kappa^N \subset \mathcal{CM}_\lambda^N$ if $0 < \kappa \leq \lambda$.
- (ii) For any $N \in \{0, \dots, 7\}$, we have $\mathcal{CM} \subset \mathcal{CM}_\lambda^N$ if $\lambda > 1$.

Remark 3.17. Furthermore, the authors show that both claims of Lemma 3.16 fail to hold for sufficiently large N , thus the following strategy cannot be extended for all $\ell \in \mathbb{N}$.

b) *Let $\ell \in \{3, 4, 5, 6, 7\}$ and $p > 2$:* Let us recall that $\tilde{G} \in \mathcal{CM}$ by Lemma 3.14. Invoking claim (ii) of Lemma 3.16 with $\lambda = p - 1 > 1$, we observe that $\tilde{G} \in \mathcal{CM}_{p-1}^\ell$. The statement then stems from the generalized Mean Value Theorem 2.15.

c) *Let $\ell = 1$ or $\ell = 2$ and $p > 1$:* For $\ell = 1$, the inequality is obvious. In fact, one straightforwardly observes that $\tilde{\mathbf{g}}^{(\ell,p)}$ always satisfies (3.2) with strict inequalities for $k = 1$, i.e.

$$(\operatorname{div}^\ell \tilde{\mathbf{g}}_n^{(\ell,p)})^{p-1} > (\operatorname{div}^\ell \tilde{\mathbf{g}}_{n+1}^{(\ell,p)})^{p-1} \quad \text{for all } n \geq 0, \quad (3.21)$$

by Lemmas 3.14 and 2.15 and Remark 3.1. For $\ell = 2$, we shall make use of the following auxiliary statement. See [24, Theorem 7] for the proof.

Lemma 3.18. *Let $N \geq 1$. Then for any $f \in \mathcal{CM}^{N+1}$, the function $f^{1/N}$ is convex.*

Let us temporarily denote $H := \operatorname{div}^2 \tilde{\mathbf{g}}^{(2,p)}$. By Lemmas 3.14 and 3.18, we have

$$\left(H^{1/N}(x)\right)' = \frac{1}{N} H^{1/N-1}(x) H'(x) < 0 \quad \text{and} \quad \left(H^{1/N}(x)\right)'' \geq 0$$

for all $N \geq 1$ and $x > 0$, therefore $H \in \mathcal{CM}_{1/N}^2$. Employing also claim (i) of Lemma 3.16, we conclude that $H \in \mathcal{CM}_\lambda^2$ for any $\lambda > 0$, hence, in particular, $H \in \mathcal{CM}_{p-1}^2$. Quick application of Lemma 2.15 reveals the desired result once again. The proof of Proposition 3.15 is complete. $\square$

In view of Theorems 2.9, Propositions 3.13 and 3.15 together imply that the sequences $\tilde{\rho}^{(\ell,p)}$, defined by (3.1), constitute p -Hardy weights of order ℓ , which are non-attainable in $\mathcal{H}_0^\ell$, for all combinations of $\ell \in \mathbb{N}$ and $p > 1$ specified in (i), (ii), and (iii).

In fact, a multitude of numerical tests using a suitable CAS such as Wolfram Mathematica suggest that not only $\tilde{G} \in \mathcal{CM}$, but also $\tilde{G} \in \mathcal{LCM}$. Consequently, we expect the following conjecture to hold true.

Conjecture 3.19. *Let $\ell \in \mathbb{N}$ and $p > 1$. The sequence $\tilde{\rho}^{(\ell,p)}$, given by (3.1), i.e.*

$$\tilde{\rho}_n^{(\ell,p)} = \frac{-\Delta_p^{(\ell)} \tilde{\mathbf{g}}_n^{(\ell,p)}}{(\tilde{\mathbf{g}}_n^{(\ell,p)})^{p-1}}$$

for all $n \geq \ell$, where $\tilde{\mathbf{g}}^{(\ell,p)}$ is defined by (2.9), is an optimal discrete p -Hardy weight of order ℓ .

Remark 3.20. In the cases where $\ell = 1$ or $p \geq 2$ is an integer, the hypothesis is an immediate consequence of Theorem 2.10 and the strict inequalities (3.20) and (3.21). In particular, for $\ell = 1$, we extended the first-order p -Hardy inequality (7), originally derived by Fisher–Keller–Pogorzelski in [13], to complex-valued sequences and, in addition, we established that the weight $\rho^{\text{FKP}} = \tilde{\rho}^{(1,p)}$ is not only null-critical, but also optimal in the sense of Definition 2.1.

4. CLASSICAL p -BIRMAN INEQUALITY

In this section, we derive a discrete analog of the classical p -Birman inequality (10) on the half-line. Specifically, we establish the following theorem, which we regard as the third main result of this work.

Theorem 4.1 (classical discrete p -Birman inequality). *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the inequality*

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p \geq \left(\frac{p-1}{p}\right)_\ell^p \sum_{n=\ell}^{\infty} \frac{|u_n|^p}{n^{\ell p}} \quad (4.1)$$

holds for all $u \in \mathcal{H}_0^\ell$ and the constant appearing on the right-hand side is optimal.

Remark 4.2. Inequality (4.1) can be equivalently written in discrete integral form. By solving the difference equation $\nabla^\ell u_n = v_n$ for $n \geq \ell$, with the initial condition $u_n = 0$ for $n < \ell$, one finds

$$u_n = \sum_{\ell \leq k_1 \leq \dots \leq k_\ell \leq n} v_{k_1} = \sum_{j=\ell}^n \binom{n+\ell-1-j}{\ell-1} v_j$$

for $n \geq \ell$. Substituting into (4.1) yields the inequality

$$\sum_{n=\ell}^{\infty} |v_n|^p \geq \left(\frac{p-1}{p}\right)_\ell^p \sum_{n=\ell}^{\infty} \left| \frac{1}{n^\ell} \sum_{k=\ell}^n \binom{n+\ell-1-k}{\ell-1} v_k \right|^p \quad (4.2)$$

for any $v \in C_0(\mathbb{N})$. In fact, using Fatou's lemma, inequality (4.2) extends to all complex sequences v . After shifting the index, Theorem 4.1 states that the generalized discrete Hardy operator

$$\mathbf{H}^{(\ell)} v_n := \frac{1}{(n+\ell-1)^\ell} \sum_{k=1}^n \binom{n+\ell-1-k}{\ell-1} v_k$$

is bounded on $\ell^p(\mathbb{N})$, for all $1 < p < \infty$, with the operator norm equal to $1/(1-1/p)_\ell$.

The classical discrete p -Birman inequality (4.1) will be determined using two distinct methodological approaches:

- First, we derive a lower bound for the optimal weights $\rho^{(\ell,p)}$, defined in (3.1), and demonstrate that these weights yield an improvement over the corresponding classical weights on the right-hand side of (4.1).
- Second, we establish a Copson-type inequality with power weights of the form n^σ . This inequality is then applied iteratively in order to obtain the classical higher-order p -Hardy inequalities.

Finally, in Section 4.3, we derive the continuous p -Birman inequality (10) from its discrete counterpart and establish the optimality of the p -Birman constant

$$B_p^{(\ell)} := \left(\frac{p-1}{p} \right)_\ell^p, \quad (4.3)$$

demonstrating that this value remains sharp also in the continuous setting. The Sections 4.2 and 4.3 closely follow our recently published paper [33], which extends the framework of [18] to the general case $p > 1$.

4.1. Lower bound for the optimal weight $\rho^{(\ell,p)}$. By detailed analysis of the optimal p -Hardy weights of order ℓ

$$\rho_n^{(\ell,p)} = \frac{-\Delta_p^{(\ell)} \mathfrak{g}_n^{(\ell,p)}}{(\mathfrak{g}_n^{(\ell,p)})^{p-1}} = \left(\frac{1}{q} \right)_\ell^{p-1} \left[\frac{\Gamma(n-\ell+1)}{\Gamma(n+1/q)} \right]^{p-1} \sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \left[\frac{\Gamma(n-\ell+j+1/q)}{\Gamma(n-\ell+j+1)} \right]^{p-1}, \quad (4.4)$$

we will prove our next principal result:

Theorem 4.3. *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the sequence $\rho^{(\ell,p)}$, given by formula (4.4), satisfies the inequality*

$$\rho_n^{(\ell,p)} > \left(\frac{p-1}{p} \right)_\ell^p \frac{1}{n^{\ell p}} \quad \text{for all } n \geq \ell.$$

Proof. The proof will be carried out separately for the cases $\ell = 1$ and $\ell \geq 2$, because Lemma 4.5, which plays a pivotal role in the latter case, does not remain valid in the regime $\ell = 1$.

a) *The case $\ell = 1$:* Since for $\ell = 1$ the weight $\rho^{(1,p)}$ assumes a relatively simple form, we establish an even stronger result: namely, that this weight admits a convergent series expansion in negative powers of n with all coefficients strictly positive, and with the leading coefficient coinciding with the classical p -Hardy constant. By formula (3.13), we have

$$\rho_n^{(1,p)} = \left(\frac{1}{q} \right)^{p-1} \left[\frac{1}{(n-1/p)^{p-1}} - \frac{1}{n^{p-1}} \right] = \left(\frac{1}{q} \right)^{p-1} n^{1-p} \left[\left(1 - \frac{1}{pn} \right)^{1-p} - 1 \right]$$

for all $n \geq 1$. Moreover, the generalized binomial theorem yields the series representation

$$\left(1 - \frac{1}{pn} \right)^{1-p} - 1 = \sum_{k=1}^{\infty} \binom{1-p}{k} \frac{(-1)^k}{p^k} \frac{1}{n^k},$$

which is valid for any $n \geq 1$, given that $p > 1$. Taking also into account the elementary identity

$$\binom{1-p}{k+1} (-1)^{k+1} = \frac{(p-1)(p)_k}{(k+1)!} \quad \text{for all } k \geq 0,$$

after shifting the index k , we finally obtain

$$\rho_n^{(1,p)} = \left(\frac{p-1}{p} \right)^p \sum_{k=0}^{\infty} \frac{(p)_k}{p^k} \frac{1}{(k+1)!} \frac{1}{n^{p+k}} > \left(\frac{p-1}{p} \right)^p \frac{1}{n^p} \quad (4.5)$$

for all $n \geq 1$.

b) *The case $\ell \geq 2$:* We now estimate the two factors appearing in (4.4) separately, by means of the following two lemmas.

Lemma 4.4. *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the inequality*

$$\frac{\Gamma(n-\ell+1)}{\Gamma(n+1/q)} > \frac{1}{n^\ell} n^{1/p}$$

holds for all $n \geq \ell$.

Proof (of Lemma 4.4). By using the recurrence relation $\Gamma(n+1) = n\Gamma(n)$, which holds for all $n \geq 1$, iteratively ℓ times, we obtain

$$\frac{\Gamma(n-\ell+1)}{\Gamma(n+1/q)} = \frac{1}{n(n-1)\cdots(n-\ell+1)} \frac{\Gamma(n+1)}{\Gamma(n+1/q)} > \frac{1}{n^\ell} \frac{\Gamma(n+1)}{\Gamma(n+1/q)}.$$

Application of the Gautschi inequality [9, Eq. (5.6.4)]

$$x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (x+1)^{1-s} \quad \text{for all } x > 0 \text{ and } s \in (0, 1), \quad (4.6)$$

with $x = n \geq \ell > 0$ and $s = 1/q \in (0, 1)$, to the last term, yields the claim of Lemma 4.4. $\square$

Lemma 4.5. *Let $\ell \geq 2$ and $p > 1$. Then the inequality*

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \left[\frac{\Gamma(n-\ell+j+1/q)}{\Gamma(n-\ell+j+1)} \right]^{p-1} \geq \left(\frac{1}{q} \right)_{\ell} \frac{1}{n^{\ell+1/q}}$$

holds for all $n \geq \ell$.

Proof (of Lemma 4.5). The proof will be divided into two steps. First, we replace the ratio of Gamma functions on the left-hand side of the inequality with a more tractable power sequence. Second, we obtain a lower bound for this new sequence that coincides with the right-hand side of the inequality above.

Step 1: We derive the inequality

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \left[\frac{\Gamma(n-\ell+j+1/q)}{\Gamma(n-\ell+j+1)} \right]^{p-1} \geq \sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \frac{1}{(n-\ell+j+1)^{1/q}} \quad (4.7)$$

for all $n \geq \ell \geq 0$, or, equivalently

$$(-1)^{\ell} \operatorname{div}^{\ell} \left[\left(\frac{\Gamma(n+1/q)}{\Gamma(n+1)} \right)^{p-1} - \frac{1}{(n+1)^{1/q}} \right] \geq 0 \quad \text{for all } \ell \geq 0 \text{ and } n \geq 0.$$

Recalling the definition of complete monotonicity (3.3), together with the generalized Mean Value Theorem 2.15, the latter inequalities will be established once we show that the function

$$F : (0, \infty) \rightarrow \mathbb{R} : x \mapsto \left(\frac{\Gamma(x+1/q)}{\Gamma(x+1)} \right)^{p-1} - \frac{1}{(x+1)^{1/q}}$$

is completely monotone, which shall be done in the remainder of this step.

First, we demonstrate that the function

$$g : (0, \infty) \rightarrow (0, \infty) : x \mapsto (x+1)^{1/p} \frac{\Gamma(x+1/q)}{\Gamma(x+1)}$$

is logarithmically completely monotone, refining the result of Proposition 3.7; see (3.6) for the corresponding definition. Observe that the function g is indeed strictly positive and temporarily denote the auxiliary function

$$h(x) = -(\log g(x))' = \psi(x+1) - \psi(x+1/q) - \frac{1}{p(x+1)} \quad \text{for all } x > 0.$$

Invoking the integral representations (3.11) for the digamma function ψ and

$$\frac{1}{x+1} = \int_0^{\infty} e^{-(x+1)t} dt \quad \text{for all } x > 0,$$

we write the function h as a Laplace transform (3.4) of the form

$$h(x) = \int_0^{\infty} e^{-xt} k(t) dt \quad \text{for all } x > 0, \quad \text{with} \quad k(t) := e^{-t} \left(\frac{e^{t/p} - 1}{1 - e^{-t}} - \frac{1}{p} \right) \quad \text{for all } t > 0.$$

Elementary computations reveal that the density k is strictly positive on the interval $(0, \infty)$, from which we infer $h \in \mathcal{CM}$, and therefore $g \in \mathcal{LCM}$.

From the Gautschi inequality (4.6), we additionally deduce that $g(x) > 1$ for all $x > 0$. Therefore, in view of Proposition 3.3, which asserts that logarithmically completely monotone functions are infinitely divisible, it follows that $g^{p-1} - 1 \in \mathcal{CM}$. Finally, multiplying this function by the completely monotone function $x \mapsto (x+1)^{-1/q}$, we conclude that $F \in \mathcal{CM}$ as well, by Proposition 3.2.

Step 2: We further estimate the right-hand side of (4.7) as

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \frac{1}{(n-\ell+j+1)^{1/q}} \geq \left(\frac{1}{q}\right)_{\ell} \frac{1}{n^{\ell+1/q}} \quad (4.8)$$

for all $n \geq \ell \geq 2$ (note that the inequality fails for $\ell = 1$). To this end, we will make use of the integral formula

$$\int_0^{\infty} t^{\lambda-1} e^{-xt} dt = \frac{\Gamma(\lambda)}{x^{\lambda}}, \quad \text{valid for all } \lambda, x > 0, \quad (4.9)$$

and rewrite the left-hand side of the latter inequality as

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \frac{1}{(n-\ell+j+1)^{1/q}} = \frac{1}{\Gamma(1/q)} \sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \int_0^{\infty} t^{1/q-1} e^{-(n-\ell+j+1)t} dt.$$

Moreover, upon interchanging the order of summation and integration, an application of the binomial theorem reveals

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \int_0^{\infty} t^{1/q-1} e^{-(n-\ell+j+1)t} dt = \int_0^{\infty} t^{1/q-1} (1-e^{-t})^{\ell} e^{(\ell-1)t} e^{-nt} dt.$$

Next, we estimate the integrand

$$(1-e^{-t})^{\ell} e^{(\ell-1)t} \geq t^{\ell}, \quad \text{or equivalently} \quad \left(\frac{e^t-1}{t}\right)^{\ell} \geq e^t.$$

Since $(e^t-1)/t \geq 1$ for all $t > 0$ and $\ell \geq 2$, it suffices to verify that

$$\left(\frac{e^t-1}{t}\right)^2 \geq e^t,$$

which, after taking square roots, stems from the elementary inequality $\sinh t/2 \geq t/2$ for all $t > 0$.

Subsequently, invoking the identity (4.9) once more, the above integral is bounded from below as

$$\int_0^{\infty} t^{1/q-1} (1-e^{-t})^{\ell} e^{(\ell-1)t} e^{-nt} dt \geq \int_0^{\infty} t^{\ell+1/q-1} e^{-nt} dt = \frac{\Gamma(\ell+1/q)}{n^{\ell+1/q}},$$

which, in turn, implies the desired inequality

$$\sum_{j=0}^{\ell} (-1)^j \binom{\ell}{j} \frac{1}{(n-\ell+j+1)^{1/q}} \geq \frac{\Gamma(\ell+1/q)}{\Gamma(1/q)} \frac{1}{n^{\ell+1/q}} = \left(\frac{1}{q}\right)_{\ell} \frac{1}{n^{\ell+1/q}}$$

for all $n \geq \ell \geq 2$. Combining inequalities (4.7) and (4.8) concludes the proof of Lemma 4.5. $\square$

Finally, from Lemmas 4.4 and 4.5, we obtain the desired estimate

$$\rho_n^{(\ell,p)} > \left(\frac{1}{q}\right)_{\ell}^{p-1} \frac{1}{n^{(p-1)\ell}} n^{1/q} \left(\frac{1}{q}\right)_{\ell} \frac{1}{n^{\ell+1/q}} = \left(\frac{p-1}{p}\right)_{\ell}^p \frac{1}{n^{\ell p}}$$

for all $n \geq \ell$. The proof of part b), and therefore of Theorem 4.3 is complete. $\square$

As an immediate corollary, we are now able to prove Theorem 4.1.

Proof (of Theorem 4.1). Since, by Theorem 3.10, the sequences $\rho^{(\ell,p)}$ constitute (optimal) p -Hardy weights of order ℓ , Theorem 4.3 yields, in turn, the classical p -Birman inequality

$$\sum_{n=\ell}^{\infty} |\nabla^{\ell} u_n|^p \geq \sum_{n=\ell}^{\infty} \rho_n^{(\ell,p)} |u_n|^p \geq \left(\frac{p-1}{p}\right)_{\ell}^p \sum_{n=\ell}^{\infty} \frac{|u_n|^p}{n^{\ell p}}$$

for all $u \in \mathcal{H}_0^{\ell}$. Notice that this inequality is sharp for every nontrivial sequence $u \neq 0$. Moreover, it is straightforward to verify that

$$\rho_n^{(\ell,p)} \sim \left(\frac{p-1}{p}\right)_{\ell}^p \frac{1}{n^{\ell p}}, \quad \text{as } n \rightarrow \infty.$$

Therefore, the optimality near infinity of the weights $\rho^{(\ell,p)}$ implies the optimality of the p -Birman constant on the right-hand side. The proof of Theorem 4.1 is complete. $\square$

In [32], we investigated another remarkable property of the optimal Hardy weights $\rho^{(\ell)}$ of order ℓ , defined in (9). In analogy with the expansion (4.5) of the weight $\rho^{(1,p)}$, we proved that $\rho_n^{(\ell)}$ admits a convergent series representation in negative powers of n , whose coefficients are all non-negative (and, in fact, strictly positive for $\ell \geq 2$) for every $n \geq \ell$. This representation is particularly useful, since including additional terms of the truncated series yields progressively tighter Birman inequalities with very explicit weights.

We observed the same behavior also for the optimal first-order p -Hardy weight $\rho^{\text{FKP}} = \tilde{\rho}^{(1,p)}$ introduced by Fisher, Keller, and Pogorzelski; see definitions (8) and (3.1). This phenomenon was originally conjectured in [13] and subsequently established in [31] using the theory of Herglotz–Nevanlinna functions. In contrast, the proof of the expansion (4.5) for the new weight $\rho^{(1,p)}$ is considerably more elementary.

For higher-order weights, the situation becomes substantially more intricate. Nevertheless, numerical experiments indicate that the following conjecture is valid. We will establish this conjecture in the cases $\ell = 1$ or $p = 2$.

Conjecture 4.6. *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the sequence $\rho^{(\ell,p)}$, given by formula (3.13), admits a convergent series expansion*

$$\rho_n^{(\ell,p)} = \left(\frac{p-1}{p}\right)_\ell \sum_{k=0}^{\infty} \frac{A_k^{(\ell,p)}}{n^{\ell p+k}} \quad \text{for all } n \geq \ell,$$

with $A_k^{(\ell,p)} > 0$ for all $k \geq 0$ and $A_0^{(\ell,p)} = 1$.

Proof (for $\ell = 1$ or $p = 2$). For $\ell = 1$, the expansion is given by formula (4.5), with

$$A_k^{(1,p)} = \frac{(p)_k}{p^k} \frac{1}{(k+1)!} \quad \text{for all } k \geq 0.$$

On the other hand, in the special case $p = 2$, the weight $\rho^{(\ell,2)}$ reduces to the expression given in (3.14). Consequently, invoking the identity [9, Eq. (26.8.11)], we have

$$\rho^{(\ell,2)} = \left(\frac{1}{2}\right)_\ell \sum_{k=0}^{\infty} \frac{S(2\ell-1+k, 2\ell-1)}{2^k} \frac{1}{n^{2\ell+k}}, \quad \text{i.e.} \quad A_k^{(\ell,2)} = \frac{S(2\ell-1+k, 2\ell-1)}{2^k}$$

for all $k \geq 0$, where

$$S(n, k) := \sum_{\substack{j_1, \dots, j_k \geq 0 \\ j_1 + \dots + j_k = n-k}} 1^{j_1} 2^{j_2} \dots k^{j_k}, \quad \text{for all } n \in \mathbb{N} \text{ and } k \in \{0, \dots, n-1\},$$

denote the *Stirling numbers of the second kind*, with the convention $S(n, n) = 1$ for all $n \in \mathbb{N}$. □

4.2. The Copson inequality with negative exponents. Here, we prove the classical p -Birman inequality (4.1) by an iterative application of a weighted discrete p -Hardy inequality, similar to the classical *Copson inequality*

$$\sum_{n=1}^{\infty} n^\sigma |\nabla u_n|^p \geq \left(\frac{p-\sigma-1}{p}\right)^p \sum_{n=1}^{\infty} n^{\sigma-p} |u_n|^p, \quad (4.10)$$

where $u \in \mathcal{H}_0^1$ and $0 \leq \sigma < p-1$. The non-negativity of σ in (4.10) is essential, as the inequality fails for $\sigma < 0$, which is precisely the regime required for our proof of (4.1). If $\sigma < 0$, a slightly smaller weight is required on the right-hand side. We present this version of Copson inequality in the next theorem, as it may be of independent interest. Copson's original work [4] contains a proof of (4.10), while the optimality of the *Copson constant*

$$C_p^{(\sigma)} := \left(\frac{p-\sigma-1}{p}\right)^p \quad (4.11)$$

is established, for example, in [2, Corollary 3].

Theorem 4.7. *Let $\sigma < 0$ and $p > 1$. Then the inequality*

$$\sum_{n=1}^{\infty} n^\sigma |\nabla u_n|^p \geq \left(\frac{p-\sigma-1}{p}\right)^p \sum_{n=1}^{\infty} (n+1)^{\sigma-p} |u_n|^p, \quad (4.12)$$

holds for all $u \in \mathcal{H}_0^1$ and the constant appearing on the right-hand side is optimal.

Proof. We prove (4.12) by choosing V_n in Theorem 2.3 as a power sequence and $\mathfrak{g}_n$, once again, as a suitable Gamma ratio. In particular, we set

$$V_n := \begin{cases} 0 & \text{if } n = 1, \\ (n-1)^\sigma & \text{if } n \geq 2, \end{cases} \quad \mathfrak{g}_n := \begin{cases} 0 & \text{if } n = 0, \\ \Gamma(n+1 - \frac{\sigma+1}{p}) / \Gamma(n) & \text{if } n \geq 1. \end{cases} \quad (4.13)$$

Since for all $n \geq 2$, we have

$$\mathfrak{g}_n - \mathfrak{g}_{n-1} = \frac{p - \sigma - 1}{p} \frac{\Gamma(n - \frac{\sigma+1}{p})}{\Gamma(n)} > 0,$$

the sequence $\mathfrak{g}$ is increasing. As the other assumptions of Theorem 2.3 are fulfilled trivially, we obtain the inequality

$$\sum_{n=2}^{\infty} (n-1)^\sigma |\nabla u_n|^p \geq \sum_{n=2}^{\infty} \rho_n |u_n|^p$$

for all $u \in \mathcal{H}_0^2$ (here we additionally require that $u_1 = 0$), where

$$\rho_n := -\frac{\operatorname{div}(V(\nabla \mathfrak{g})^{p-1})_n}{(\mathfrak{g}_n)^{p-1}} = \left(\frac{p - \sigma - 1}{p} \right)^{p-1} n^{\sigma-p+1} H\left(\frac{1}{n}\right) \quad (4.14)$$

and H is an auxiliary function defined on the interval $[0, 1)$ by

$$H(x) := \frac{(1-x)^\sigma}{(1 - \frac{\sigma+1}{p}x)^{p-1}} - 1.$$

In order to establish the inequality (4.12), it suffices to show that H satisfies

$$H(x) > \frac{p - \sigma - 1}{p} x \quad \text{for all } x \in (0, 1). \quad (4.15)$$

Since $H(0) = 0$ and $H'(0) = (p - \sigma - 1)/p$, inequality (4.15) can be written as $H(x) > H(0) + H'(0)x$, and so it holds true if H is strictly convex on $(0, 1)$. A routine calculation yields

$$H''(x) = \frac{p - \sigma - 1}{p^2} \frac{(1-x)^{\sigma-2}}{(1 - \frac{\sigma+1}{p}x)^{p+1}} \left[(p - \sigma)(1 + \sigma)^2 x^2 - 2(1 + \sigma)(p - \sigma)x + p(1 - \sigma) \right].$$

Since the discriminant $4\sigma(1 + \sigma)^2(p - 1)(p - \sigma)$ of the quadratic term is negative, we conclude that $H''(x) > 0$ for all $x \in (0, 1)$, indeed.

Furthermore, in complete analogy with the proof of Theorem 2.10 in Section 2.1.4, one can verify that the weight ρ , defined by (4.14), is optimal (the notion of optimality given in Definition 2.1 can be extended in a straightforward manner to the setting of weighted Hardy inequalities). In particular, ρ is optimal near infinity, and hence the optimality of the Copson constant (4.11) on the right-hand side of inequality (4.12) follows immediately from the asymptotics

$$\rho_n \sim \left(\frac{p - \sigma - 1}{p} \right)^p n^{\sigma-p}, \quad \text{as } n \rightarrow \infty.$$

The details are left as an exercise for the reader. See also Remark 4.13 for an alternative proof of the optimality of the Copson constant. The proof of Theorem 4.7 is complete. $\square$

Remark 4.8. The Copson inequality (4.10) can also be established using the approach described above. Assuming $0 \leq \sigma < p - 1$, we apply Theorem 2.3 with $V_n := n^\sigma$ and $\mathfrak{g}_n$ defined in (4.13). This yields the inequality

$$\sum_{n=1}^{\infty} n^\sigma |\nabla u_n|^p \geq \sum_{n=1}^{\infty} \rho_n |u_n|^p, \quad (4.16)$$

with ρ_n given by (4.14), but with the function H redefined on $[0, 1]$ by

$$H(x) := \left(1 - \frac{\sigma + 1}{p} x \right)^{1-p} - (1 + x)^\sigma.$$

Although this version of H is no longer convex on $(0, 1)$, inequality (4.15) remains true and can be verified as follows. First, we observe that $H^{(\lceil \sigma \rceil + 1)}(x) > 0$ for all $x \in (0, 1)$. Noticing also that $H(0) = 0$ and

$H'(0) = (p - \sigma - 1)/p$, the Taylor–Lagrange formula implies, for any $x \in (0, 1)$, the existence of $\xi \in (0, 1)$ such that

$$H(x) = \frac{p - \sigma - 1}{p}x + \sum_{k=2}^{\lceil \sigma \rceil} \frac{H^{(k)}(0)}{k!}x^k + \frac{H^{(\lceil \sigma \rceil + 1)}(\xi)}{(\lceil \sigma \rceil + 1)!}x^{\lceil \sigma \rceil + 1},$$

from which the desired inequality follows if $H^{(k)}(0) \geq 0$ for all $k \in \{2, \dots, \lceil \sigma \rceil\}$. To verify also the last claim and finish the proof, we estimate

$$\begin{aligned} H^{(k)}(0) &= \frac{(p + k - 2)(p + k - 3) \cdots (p - 1)}{p^k} (\sigma + 1)^k - \sigma(\sigma - 1) \cdots (\sigma - k + 1) \\ &\geq \left(1 - \frac{1}{p}\right) (\sigma + 1)^k - \sigma^k \geq \left(1 - \frac{1}{\sigma + 1}\right) (\sigma + 1)^k - \sigma^k = \sigma(\sigma + 1)^{k-1} - \sigma^k \geq 0. \end{aligned}$$

Furthermore, inequality (4.16), once again, offers an optimal version of the Copson inequality, which can be of independent interest, as the existing literature on such refinements, for example [5, 6, 7], seems to consider only the special case $p = 2$.

Proof (of Theorem 4.1). Here, we restrict our attention to establishing the inequality (4.1). An alternative proof of the optimality of the p -Birman constant $B_p^{(\ell)}$, distinct from the argument developed in Section 4.1 and independent of the optimal weights obtained in Theorem 3.10, is provided in Remark 4.13.

The proof proceeds by induction on $\ell \in \mathbb{N}$. For the base case $\ell = 1$, inequality (4.1) coincides with the classical p -Hardy inequality (1).

Assume $\ell \geq 2$ and pick $u \in \mathcal{H}_0^\ell$. Recalling definitions of the discrete gradient and divergence (1.5), we have the identities $\operatorname{div} = \nabla S$ and $\operatorname{div} \nabla = \nabla \operatorname{div}$. It follows that

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p = \sum_{n=\ell-1}^{\infty} |\nabla^{\ell-1} \operatorname{div} u_n|^p \geq B_p^{(\ell-1)} \sum_{n=\ell-1}^{\infty} \frac{|\operatorname{div} u_n|^p}{n^{(\ell-1)p}},$$

where we applied the induction hypothesis on the sequence $\operatorname{div} u \in \mathcal{H}_0^{\ell-1}$. Using once more the identity $\operatorname{div} = \nabla S$ and applying the already proven inequality (4.12) with $\sigma = -(\ell - 1)p < 0$ and the sequence $Su \in \mathcal{H}_0^1$, we further estimate

$$\sum_{n=\ell-1}^{\infty} \frac{|\operatorname{div} u_n|^p}{n^{(\ell-1)p}} = \sum_{n=1}^{\infty} \frac{|\nabla S u_n|^p}{n^{(\ell-1)p}} \geq C_p^{((1-\ell)p)} \sum_{n=1}^{\infty} \frac{|S u_n|^p}{(n+1)^{\ell p}}.$$

Shifting the index in the last sum and noticing that $B_p^{(\ell-1)} C_p^{((1-\ell)p)} = B_p^{(\ell)}$, see formulas (4.3) and (4.11), the above inequalities combine to

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n|^p \geq B_p^{(\ell)} \sum_{n=\ell}^{\infty} \frac{|u_n|^p}{n^{\ell p}}.$$

The proof of inequality (4.1) is complete. $\square$

4.3. Continuous p -Birman inequality. In this section, we derive the continuous p -Birman inequality from its discrete counterpart (4.1), following an approach utilized for $p = 2$ in [18]. Beyond the fact that the transition from discrete to continuous inequalities is interesting in itself, this step is motivated by two other objectives. First, it provides an alternative proof of the p -Birman inequality to the standard procedure, which naturally avoids the discrete formulation and is indicated in Remark 4.12. Second, we take advantage of this passage to prove the optimality of constants in Theorems 4.1 and 4.7 in Remark 4.13.

Theorem 4.9 (continuous p -Birman inequality). *Let $\ell \in \mathbb{N}$ and $p > 1$. Then the inequality*

$$\int_0^\infty |\varphi^{(\ell)}(x)|^p dx \geq \left(\frac{p-1}{p}\right)_\ell^p \int_0^\infty \frac{|\varphi(x)|^p}{x^{\ell p}} dx, \quad (4.17)$$

holds for all $\varphi \in C_0^\infty(\mathbb{R}_+)$ and the constant appearing on the right-hand side is optimal.

The proof will be developed progressively over the course of this section. We begin by establishing inequality (4.17), after which we demonstrate the optimality of the p -Birman constant $B_p^{(\ell)}$ in (4.17). Prior to the proof of inequality (4.17), we state a preparatory lemma that relates the discrete derivatives to the continuous ones.

Lemma 4.10. For $N > 0$, $n \in \mathbb{N}_0$, and $\varphi \in C_0^\infty(\mathbb{R}_+)$, we denote $v_n^N := \varphi(n/N)$. Then for any $\ell \in \mathbb{N}$, we have

$$\nabla^\ell v_n^N = \frac{1}{N^\ell} \varphi^{(\ell)}\left(\frac{n}{N}\right) + \mathcal{O}\left(\frac{1}{N^{\ell+1}}\right), \quad \text{as } N \rightarrow \infty,$$

uniformly for all $n \geq \ell$, i.e. the constant hidden in the Landau $\mathcal{O}$ -symbol is n -independent.

Proof. By definition (1.5), we have

$$\nabla^\ell v_n^N = \sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} v_{n-k}^N.$$

Moreover, applying the Taylor–Lagrange formula, we obtain

$$v_{n-k}^N = \varphi\left(\frac{n-k}{N}\right) = \sum_{j=0}^{\ell} \frac{1}{j!} \varphi^{(j)}\left(\frac{n}{N}\right) \left(\frac{-k}{N}\right)^j + \mathcal{O}\left(\frac{1}{N^{\ell+1}}\right),$$

where the constant in the Landau $\mathcal{O}$ -symbol can be taken n -independent since $\varphi \in C_0^\infty(\mathbb{R}_+)$. It follows that

$$\begin{aligned} \nabla^\ell v_n^N &= \sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} \left[\sum_{j=0}^{\ell} \frac{1}{j!} \varphi^{(j)}\left(\frac{n}{N}\right) \left(\frac{-k}{N}\right)^j + \mathcal{O}\left(\frac{1}{N^{\ell+1}}\right) \right] \\ &= \sum_{j=0}^{\ell} \frac{(-1)^j}{j!} \left[\sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} k^j \right] \frac{1}{N^j} \varphi^{(j)}\left(\frac{n}{N}\right) + \mathcal{O}\left(\frac{1}{N^{\ell+1}}\right), \end{aligned}$$

as $N \rightarrow \infty$, uniformly in $n \geq \ell$. Now, it suffices to apply the identity [9, Eq. (26.8.6)]

$$\sum_{k=0}^{\ell} (-1)^k \binom{\ell}{k} k^j = \begin{cases} 0 & \text{for } j < \ell, \\ (-1)^\ell \ell! & \text{for } j = \ell, \end{cases}$$

to the term in the square brackets to finish the proof. $\square$

Proof (of inequality (4.17)). First notice that, by scaling of the variable x in both sides of (4.17), it is sufficient to prove the inequality for test functions φ supported in $(0, 1)$.

Fix $\varphi \in C_0^\infty(0, 1)$, $\ell \in \mathbb{N}$, and define $v_n^N := \varphi(n/N)$ for all $n \in \mathbb{N}_0$ and $N > 0$ sufficiently large, such that $v_n^N = 0$ for $n < \ell$. Since also $v_n^N = 0$ for all $n > N$, applying Lemma 4.10, we infer that

$$\sum_{n=\ell}^{\infty} |\nabla^\ell v_n^N|^p = \sum_{n=1}^N |\nabla^\ell v_n^N|^p = \sum_{n=1}^N \left| \frac{1}{N^\ell} \varphi^{(\ell)}\left(\frac{n}{N}\right) + \mathcal{O}\left(\frac{1}{N^{\ell+1}}\right) \right|^p = \frac{1}{N^{\ell p}} \sum_{n=1}^N \left| \varphi^{(\ell)}\left(\frac{n}{N}\right) \right|^p + \mathcal{O}\left(\frac{1}{N^{\ell p}}\right),$$

as $N \rightarrow \infty$. For the sequence $u^N := N^{\ell-1/p} v^N$ it follows that

$$\sum_{n=\ell}^{\infty} |\nabla^\ell u_n^N|^p = \frac{1}{N} \sum_{n=1}^N \left| \varphi^{(\ell)}\left(\frac{n}{N}\right) \right|^p + \mathcal{O}\left(\frac{1}{N}\right), \quad \text{as } N \rightarrow \infty.$$

Since also $u_n^N = 0$ for $n < \ell$ and $n > N$, we may apply the discrete p -Birman inequality (4.1) to obtain

$$\sum_{n=1}^N \frac{1}{N} \left| \varphi^{(\ell)}\left(\frac{n}{N}\right) \right|^p + \mathcal{O}\left(\frac{1}{N}\right) \geq \left(\frac{p-1}{p}\right)_\ell^p \sum_{n=\ell}^{\infty} \frac{|u_n^N|^p}{n^{\ell p}} = \left(\frac{p-1}{p}\right)_\ell^p \sum_{n=1}^N \frac{1}{N} \frac{|\varphi(n/N)|^p}{(n/N)^{\ell p}}$$

for $N \rightarrow \infty$. Recognizing the Riemann sums and taking the limit $N \rightarrow \infty$ in the above inequality, we arrive at the desired integral inequality (4.17). $\square$

Remark 4.11. We remark, without going into details, that one can proceed analogously as in the proof of inequality (4.17), using, this time, the sampling sequence

$$u_n^N := N^{1-(\sigma+1)/p} \varphi\left(\frac{n}{N}\right),$$

to deduce from (4.10) and (4.12) their continuous counterpart

$$\int_0^\infty x^\sigma |\varphi'(x)|^p dx \geq \left(\frac{p-\sigma-1}{p}\right)^p \int_0^\infty x^{\sigma-p} |\varphi(x)|^p dx, \quad (4.18)$$

which holds for all $\varphi \in C_0^\infty(\mathbb{R}_+)$ and $\sigma < p-1$ (the continuous limit takes the same form regardless of the sign of σ).

Remark 4.12. An integral form of the inequality (4.18) appeared already in the classical book by Hardy, Littlewood, and Pólya [17, Theorem 330], see also [23, Theorem 2]. The continuous p -Birman inequality (4.17) can be quickly proved through an iterative application of (4.18). Specifically, starting with the ℓ -th derivative of a test function φ and selecting $\sigma = -(j-1)p$ in the j -th step for $j \in \{1, \dots, \ell\}$, the result follows directly.

Proof (of optimality of the p -Birman constant $B_p^{(\ell)}$ in Theorem 4.9). We verify the optimality of the constant $B_p^{(\ell)}$ in the continuous p -Birman inequality (4.17), which amounts to proving

$$\inf_{0 \neq \varphi \in C_0^\infty(\mathbb{R}_+)} \frac{\int_0^\infty |\varphi^{(\ell)}(x)|^p dx}{\int_0^\infty |\varphi(x)|^p / x^{\ell p} dx} = B_p^{(\ell)} \quad (4.19)$$

by construction of an approximating sequence of test functions $\varphi_N \in C_0^\infty(\mathbb{R}_+)$, for which the ratio in the infimum converges to $B_p^{(\ell)}$ as $N \rightarrow \infty$. Concretely for $N > 2$ and $x > 0$, we set

$$\varphi_N(x) := x^{\ell-1/p} \xi_N(x), \quad (4.20)$$

with the cut-off function $\xi_N \in C_0^\infty(\mathbb{R}_+)$ satisfying

$$\xi_N(x) = \begin{cases} 1 & \text{if } x \in [2/N, 1], \\ 0 & \text{if } x \in \mathbb{R}_+ \setminus [1/N, 2], \end{cases} \quad \text{and} \quad |\xi_N^{(j)}(x)| \leq \begin{cases} CN^j & \text{if } x \in [1/N, 2/N], \\ C & \text{if } x \in [1, 2], \end{cases} \quad (4.21)$$

for all $j \in \{0, \dots, \ell\}$, where $C > 0$ is a constant. Such a function can be constructed, for instance, as

$$\xi_N(x) := \eta(Nx - 1)\eta(2 - x),$$

where η is a smooth function on $\mathbb{R}$ satisfying $\eta(x) = 0$ for $x \leq 0$ and $\eta(x) = 1$ for $x \geq 1$. This regularization has been utilized in a slightly different situation in [25, Note 2.6].

Using the Leibniz rule, we estimate the L^p -norm of the ℓ -th derivative of φ_N as

$$\left(\int_0^\infty |\varphi_N^{(\ell)}(x)|^p dx \right)^{1/p} \leq \sum_{j=0}^{\ell} \binom{\ell}{j} \left(1 - \frac{1}{p} + j \right)_{\ell-j} \left(\int_0^\infty |x^{j-1/p} \xi_N^{(j)}(x)|^p dx \right)^{1/p}.$$

Employing the properties (4.21), we further estimate the integrals in the last sum separately for $j = 0$ and $j \in \{1, \dots, \ell\}$. For $j = 0$, we have

$$\int_0^\infty |x^{-1/p} \xi_N(x)|^p dx \leq C^p \int_{1/N}^{2/N} \frac{dx}{x} + \int_{2/N}^1 \frac{dx}{x} + C^p \int_1^2 \frac{dx}{x} = 2C^p \log 2 + \log(N/2),$$

while for $j \in \{1, \dots, \ell\}$, we obtain

$$\int_0^\infty |x^{j-1/p} \xi_N^{(j)}(x)|^p dx \leq C^p N^{jp} \int_{1/N}^{2/N} x^{jp-1} dx + C^p \int_1^2 x^{jp-1} dx = 2C^p \frac{2^{jp} - 1}{jp}.$$

Recalling formula (4.3), we obtain the upper bound

$$\left(\int_0^\infty |\varphi_N^{(\ell)}(x)|^p dx \right)^{1/p} \leq \left[B_p^{(\ell)} (A_0 + \log(N/2)) \right]^{1/p} + A_1,$$

where A_0 and A_1 are N -independent constants. Conversely, we find

$$\int_0^\infty \frac{|\varphi_N(x)|^p}{x^{\ell p}} dx \geq \int_{2/N}^1 \frac{dx}{x} = \log(N/2).$$

Altogether, bearing the continuous p -Birman inequality (4.17) in mind, we have

$$B_p^{(\ell)} \leq \frac{\int_0^\infty |\varphi_N^{(\ell)}(x)|^p dx}{\int_0^\infty |\varphi_N(x)|^p / x^{\ell p} dx} \leq \frac{\left(\left[B_p^{(\ell)} (A_0 + \log(N/2)) \right]^{1/p} + A_1 \right)^p}{\log(N/2)} \longrightarrow B_p^{(\ell)}, \quad \text{as } N \rightarrow \infty,$$

which yields (4.19). The proof of Theorem 4.9 is complete. $\square$

Remark 4.13. The proof of the continuous p -Birman inequality shows that inequality (4.17) follows from (4.1) regardless of the explicit form of the constant $B_p^{(\ell)}$. Consequently, the optimality of the constant $B_p^{(\ell)}$ defined by (4.3) in the continuous p -Birman inequality (4.17) immediately implies its optimality in the discrete counterpart (4.1).

Moreover, the proof of the optimality of the Copson constant $C_p^{(\sigma)}$ in the continuous inequality (4.18), and consequently in the discrete counterparts (4.10) and (4.12), follows an argument similar to the one detailed above. In this case, the test functions (4.20) are replaced by

$$\varphi_N(x) := x^{1-(\sigma+1)/p} \xi_N(x).$$

The sharpness within a space of non-smooth functions is also proved in [23, Theorem 2(iii)].

REFERENCES

- [1] Alzer, H. On some inequalities for the Gamma and Psi functions. *Math. Comp.* 66, 217 (1997), 373–389.
- [2] Bennett, G. Some elementary inequalities. *Quart. J. Math.* 38, 4 (1987), 401–425.
- [3] Birman, M. Sh. On the spectrum of singular boundary-value problems. *Mat. Sb. (N.S.)* 55, 97 (1961), 125–174.
- [4] Copson, E. T. Note on Series of Positive Terms. *J. Lond. Math. Soc. s1-3*, 1 (1928), 49–51.
- [5] Das, B., and Manna, A. On the improvements of Hardy and Copson inequalities. *Rev. Real Acad. Cienc. Exactas Fis. Nat. Ser. A-Mat.* 117, 2 (2023), art. no. 92.
- [6] Das, B., and Manna, A. An improved Copson inequality. arXiv:2508.00388 [math.CA], (2025).
- [7] Das, B., Manna, A., and Paul, T. Improvement of the discrete weighted variant Hardy’s inequality and criticality of the improved weight. *J. Geom. Anal.* 35, 12 (2025), art. no. 376.
- [8] Davies, E., and Hinz, A. Explicit constants for Rellich inequalities in $L_p(\Omega)$. *Math. Z.* 227, 3 (1998), 511–523.
- [9] *NIST Digital Library of Mathematical Functions*. <https://dlmf.nist.gov/>, Release 1.1.5 of 2022-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller, B. V. Saunders, H. S. Cohl, and M. A. McClain, eds.
- [10] Fischer, F. A non-local quasi-linear ground state representation and criticality theory. *Calc. Var. Partial Differential Equations* 62, 5 (2023), art. no. 163.
- [11] Fischer, F. *Hardy Inequalities on Graphs*. Doctoral dissertation, Universität Potsdam, 2024.
- [12] Fischer, F. On the optimality and decay of p -Hardy weights on graphs. *Calc. Var. Partial Differential Equations* 63, 7 (2024), art. no. 162.
- [13] Fischer, F., Keller, M., and Pogorzelski, F. An Improved Discrete p -Hardy Inequality. *Integr. Equ. Oper. Theory* 95, 1 (2023), art. no. 24.
- [14] Gerhat, B., Krejčířík, D., and Štampach, F. Criticality transition for positive powers of the discrete Laplacian on the half line. *Rev. Mat. Iberoam.* 41, 3 (2025), 1173–1200.
- [15] Gerhat, B., Krejčířík, D., and Štampach, F. An improved discrete Rellich inequality on the half-line. *Israel J. Math.* 268, 1 (2025), 45–72.
- [16] Gesztesy, F., Littlejohn, L. L., Michael, I., and Wellman, R. On Birman’s sequence of Hardy–Rellich-type inequalities. *J. Differential Equations* 264, 4 (2018), 2761–2801.
- [17] Hardy, G. H., Littlewood, J. E., and Pólya, G. *Inequalities*. Cambridge University Press, 1952.
- [18] Huang, X., and Ye, D. One-dimensional sharp discrete Hardy–Rellich inequalities. *J. Lond. Math. Soc.* 109, 1 (2024), art. no. e12838.
- [19] Keller, M., Pinchover, Y., and Pogorzelski, F. An improved discrete Hardy inequality. *Amer. Math. Monthly* 125, 4 (2018), 347–350.
- [20] Keller, M., Pinchover, Y., and Pogorzelski, F. Optimal Hardy inequalities for Schrödinger operators on graphs. *Comm. Math. Phys.* 358, 2 (2018), 767–790.
- [21] Keller, M., Pinchover, Y., and Pogorzelski, F. Criticality theory for Schrödinger operators on graphs. *J. Spectr. Theory* 10, 1 (2020), 73–114.
- [22] Kufner, A., Maligranda, L., and Persson, L.-E. The prehistory of the Hardy inequality. *Amer. Math. Monthly* 113, 8 (2006), 715–732.
- [23] Kufner, A., Maligranda, L., and Persson, L.-E. *The Hardy Inequality: About Its History and Some Related Results*. Vydavatelský Servis, 2007.
- [24] Lorch, L., and Newman, D. J. On the Composition of Completely Monotonic Functions and Completely Monotonic Sequences and Related Questions. *J. Lond. Math. Soc.* 28, 1 (1983), 31–45.
- [25] Owen, M. P. The Hardy–Rellich inequality for polyharmonic operators. *Proc. R. Soc. Edinb. A* 129, 1 (1999), 825–839.
- [26] Rellich, F. Halbbeschränkte Differentialoperatoren höherer Ordnung. In *Proceedings of the International Congress of Mathematicians, 1954, vol. 3*, North-Holland Publishing Co., pp. 243–250.
- [27] Schilling, R., Song, R., and Vondracek, Z. *Bernstein Functions: Theory and Applications*. De Gruyter, 2012.
- [28] van Haeringen, H. Completely monotonic and related functions. *J. Math. Anal. Appl.* 204, 2 (1996), 389–408.
- [29] van Haeringen, H. Inequalities for real powers of completely monotonic functions. *J. Math. Anal. Appl.* 210, 1 (1997), 102–113.
- [30] Widder, D. V. *The Laplace Transform*. Princeton University Press, 1941.
- [31] Štampach, F., and Waclawek, J. A Herglotz–Nevanlinna function from the optimal discrete p -Hardy weight. *Proc. Amer. Math. Soc.* 154, 2 (2026), 807–819.
- [32] Štampach, F., and Waclawek, J. Optimal discrete Hardy–Rellich–Birman inequalities. *Journal d’Analyse Mathématique*, (2025). Accepted for publication.
- [33] Štampach, F., and Waclawek, J. The p -Hardy–Rellich–Birman inequalities on the half-line. arXiv:2603.08864 [math.CA], (2026).

Acknowledgment. I would like to express my sincere gratitude to my supervisor, František Štampach, for his patient guidance, continuous support, and insightful discussions, which greatly contributed to the development of this work. I also thank him for his careful reading of the manuscript.

(Bc. Jakub Waclawek) DEPARTMENT OF MATHEMATICS, FACULTY OF NUCLEAR SCIENCES AND PHYSICAL ENGINEERING, CZECH TECHNICAL UNIVERSITY IN PRAGUE, TROJANOVA 13, 12000 PRAHA 2, CZECH REPUBLIC
Email address: waclajak@cvut.cz