A natural derivative on $[0, n]$ and a binomial Poincaré inequality

Erwan Hillion* Oliver Johnson† Yaming Yu‡

December 20, 2013

Abstract

We consider probability measures supported on a finite discrete interval $[0, n]$. We introduce a new finite difference operator ∇_n, defined as a linear combination of left and right finite differences. We show that this operator ∇_n plays a key role in a new Poincaré (spectral gap) inequality with respect to binomial weights, with the orthogonal Krawtchouk polynomials acting as eigenfunctions of the relevant operator. We briefly discuss the relationship of this operator to the problem of optimal transport of probability measures.

2010 Mathematics Subject Classification: 46N30 (primary); 60E15 (secondary)

1 Introduction and main results

Many results in functional analysis are better understood in the context of continuous spaces than discrete. One reason that the real-valued case is more tractable than integer-valued problems is the existence of a spatial derivative $\frac{\partial}{\partial x}$, well-defined in the sense that the left and right derivatives coincide for a large class of functions. However, the situation is more complicated for integer-valued functions f. There exist two competing derivatives ∇^l and ∇^r, defined as $\nabla^l f(k) = f(k) - f(k-1)$ and $\nabla^r f(k) = f(k+1) - f(k)$, which are adjoint with respect to counting measure on $\mathbb{Z}$. In this paper, we define a new finite difference operator for functions on $[0, n]$, which interpolates between ∇^l and ∇^r.

Definition 1.1. *Fix an integer $n \geq 1$, and denote by ∇_n the finite difference operator defined by*

$$\begin{aligned} (\nabla_n f)(k) &= \frac{k}{n}(\nabla^l f)(k) + \frac{n-k}{n}(\nabla^r f)(k) \\ &= \frac{k}{n}(f(k) - f(k-1)) + \frac{n-k}{n}(f(k+1) - f(k)). \end{aligned} \tag{1}$$

We will argue that this operator has certain desirable properties, and as such deserves further attention. In particular, we will show that in two senses it is a natural choice of derivative in relation to binomial measures $b_{n,t}(k) = \binom{n}{k} t^k (1-t)^{n-k}$. The question of the uniqueness of whether this the unique choice of derivate operator with such properties remains open, expect for the easy cases where $n = 1$ or 2.

Firstly, in Section 2, we will show that this operator ∇_n acts like the translation operator on the real line. That is, in Equation (10) below, we describe how a probability measure μ on $\mathbb{R}$ can be smoothly translated using a sequence of intermediate measures μ_t. Equation (10) describes the effect of this translation action through its effect on arbitrary test functions f. We prove the following theorem, which acts as a discrete counterpart of (10), with the relationship between measure $b_{n,t}$ and operator ∇_n playing a key role:

*Statistics Group, Department of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK. Email `Erwan.Hillion@bris.ac.uk`

†Statistics Group, Department of Mathematics, University of Bristol, University Walk, Bristol, BS8 1TW, UK. Email `o.johnson@bris.ac.uk`

‡Department of Statistics, University of California, Irvine, CA 92697, USA. Email `yamingy@uci.edu`

Theorem 1.2. *The operator ∇_n gives a smooth translation of point masses from point 0 to point n using the binomial measures $b_{n,t}$ in that*

1. *$b_{n,t}$ satisfies the initial condition $b_{n,0} = \delta_0$ and the final condition $b_{n,1} = \delta_n$.*

2. *For every function $f : \mathbb{Z} \to \mathbb{R}$,*

$$\frac{\partial}{\partial t} \sum_{k \in \mathbb{Z}} f(k) b_{n,t}(k) = n \sum_{k \in \mathbb{Z}} (\nabla_n f)(k) b_{n,t}(k). \tag{2}$$

Secondly, in Proposition 3.2 below we will show that the map ∇_n and its adjoint $\widetilde{\nabla}_n$ (with respect to binomial weights) act as ladder operators for the Krawtchouk polynomials ϕ_r (see Theorem 3.1). This allows us to describe the spectrum of the map $\left(\widetilde{\nabla}_n \circ \nabla_n\right)$, with ϕ_r being eigenfunctions with eigenvalue $\frac{r(n-r+1)}{n^2 t(1-t)}$. In particular, taking the smallest non-zero eigenvalue leads to a Poincaré (spectral gap) inequality for the binomial law, using the natural derivative operator ∇_n, and gives the case of equality.

Theorem 1.3. *Fix $t \in (0,1)$ and consider function $f : \{0, \dots n\} \to \mathbb{R}$ satisfying $\sum_{k=0}^n f(k) b_{n,t}(k) = 0$. Then*

$$\sum_{k=0}^n b_{n,t}(k) f(k)^2 \leq nt(1-t) \sum_{k=0}^n b_{n,t}(k) \left(\nabla_n f(k)\right)^2. \tag{3}$$

Equality holds if and only if f is a linear combination of $\phi_1(k) = \frac{1}{1-t}(k - nt)$ and $\phi_n(k) = n! \left(\frac{-t}{1-t}\right)^{n-k}$.

The idea of studying Poincaré inequalities with respect to discrete distributions is not a new one. For example, Bobkov and co-authors [1, 2, 3, 4] give results concerning probability measures supported on the discrete cube (with the difference ∇^r taken modulo 2). Cacoullos [5], Chen and Lou [6] and Klaasen [9] give results concerning ∇^r on $\mathbb{Z}$ and $\mathbb{Z}^n$. In particular, Table 2.1 of Klaassen [9] shows that for Poisson mass function Π_λ, if $\sum_k f(k)\Pi_\lambda(k) = 0$ then

$$\sum_{k=0}^\infty \Pi_\lambda(k) f(k)^2 \leq \lambda \sum_{k=0}^\infty \Pi_\lambda(k) \left(\nabla^r f(k)\right)^2. \tag{4}$$

This can be understood as a consequence of the fact that ∇^r (and its adjoint with respect to Poisson weights $\widetilde{\nabla}^r$) act as ladder operators with respect to Poisson-Charlier polynomials, meaning that the Poisson-Charlier polynomials are eigenfunctions of $\left(\widetilde{\nabla}^r \circ \nabla^r\right)$. These results also have an analogy with the work of Chernoff [7], where the corresponding result was proved for normal random variables, with the Hermite polynomials acting as eigenfunctions of the corresponding map.

However, Klaassen does not deduce such a clean result for binomial weights, requiring a weighting term on the right-hand side

$$\sum_{k=0}^n b_{n,t}(k) f(k)^2 \leq t \sum_{k=0}^n b_{n,t}(k)(n-k) \left(\nabla^r f(k)\right)^2 \tag{5}$$

Note that (see Remark 3.3 below) that our Theorem 1.3 is a stronger result than Klaassen's Equation (5). Note that as $n \to \infty$ with $tn = \lambda$, Theorem 1.3 converges to Equation (4).

Note that although we do not directly discuss applications here, in other settings the rate of convergence in variance of reversible Markov chains can be bounded in terms of the spectral gap (see for example [10, Lemma 2.1.4]).

In general, Poincaré inequalities are often viewed as a consequence of log-Sobolev inequalities (see for example [10, Lemma 2.2.2]). In particular, for Poisson measures Π_λ, Bobkov and Ledoux [4, Corollary 4] prove that for any positive function f,

$$\mathrm{Ent}_{\Pi_\lambda}(f) \leq \lambda \sum_{k=0}^{\infty} \Pi_\lambda(k) \frac{\left(\nabla^r f(k)\right)^2}{f(k)}, \tag{6}$$

and show that Klaasen's Poincaré inequality (4) can be deduced from (6). Here, $\mathrm{Ent}_\nu(f) = \sum_k \Theta(f(k))\nu(k) - \Theta\left(\sum_k f(k)\nu(k)\right)$, where $\Theta(t) = t\log t$. It is natural to conjecture that an equivalent of Equation (6) should hold for Binomial random variables with our natural derivative ∇_n, that is

$$\mathrm{Ent}_{b_{n,t}}(f) \leq nt(1-t) \sum_{k=0}^{n} b_{n,t}(k) \frac{\left(\nabla_n f(k)\right)^2}{f(k)}. \tag{7}$$

However, this result (7) is in general false. Consider for example $n = 2$, $t = 1/2$, $f(0) = f(2) = 9/10$, $f(1) = 1/10$. In this case, $\mathrm{Ent}_{b_{n,t}}(f) = 0.18403$ and the right-hand side of Equation (7) is 0.17777, and the inequality fails. The question of natural conditions on f under which Equation (7) holds remains open.

The structure of the remainder of the paper is as follows. In Section 2, we discuss the translation problem in $\mathbb{Z}$ and prove the existence of a fundamental solution for the problem under the choice of ∇ as the ∇_n from Definition 1.1. In Section 3 we prove Proposition 3.2, the key result leading to the Poincaré inequality Theorem 1.3.

2 The translation problem in $\mathbb{Z}$

It is clear that there exists an unambiguous definition of translations of real-valued probability measures, defined as the push-forward of the translation map. That is, let μ be a probability measure on $\mathbb{R}$ (with its Borel σ-algebra) having a smooth density ρ w.r.t. the Lebesgue measure dx. The n-translation of μ, where $n \in \mathbb{R}$, is the family of measures $(\mu_t = \rho_t dx)_{t\in[0,1]}$, where the density ρ_t is defined by

$$\forall x \in \mathbb{R},\ \rho_t(x) = \rho(x - nt). \tag{8}$$

In other words, the measure μ_t is the push-forward of μ by the translation map $T_t(x) = x + nt = (1-t)x + t(x+n)$. In particular,

$$\frac{\partial}{\partial t}\rho_t(x) = -n\frac{\partial}{\partial x}\rho_t(x). \tag{9}$$

This can be generalized for non absolutely continuous probability measures, writing Equation (9) in the sense of distributions:

$$\frac{\partial}{\partial t}\int_{\mathbb{R}} f(x) d\mu_t(x) = n\int_{\mathbb{R}} \frac{\partial}{\partial x} f(x) d\mu_t(x), \qquad \text{for all } f \in \mathcal{C}_c^\infty(\mathbb{R}). \tag{10}$$

This equation means that the measure μ_t is the convolution of the initial measure μ_0 with the fundamental solution of Equation (10):

$$\mu_t = \mu_0 * \delta(x - nt). \tag{11}$$

Notice that this construction of μ_t allows a smooth interpolation of probability measures. In this paper we generalize these heuristics to the case of probability measures on $\mathbb{Z}$.

Definition 2.1. *A probability measure μ_1 on $\mathbb{Z}$ is the n-translation of another probability measure μ_0 if*

$$\mu_1(k+n) = \mu_0(k) \qquad \textit{for all } k \in \mathbb{Z}.$$

In particular, we will consider measures that smoothly interpolate between point masses

$$\mu_0 = \delta_0 \quad \textit{and} \quad \mu_1 = \delta_n. \tag{12}$$

The non-connectedness of $\mathbb{Z}$ makes it impossible to generalize Equation (8) directly. However, we will adapt the "PDE point of view", given in Equation (10), to construct the n-translation of point masses (12), in a way that satisfies

$$\frac{\partial}{\partial t}\sum_{k\in\mathbb{Z}} f(k)\mu_t(k) = n\sum_{k\in\mathbb{Z}} \nabla f(k)\mu_t(k). \tag{13}$$

The main problem in this adaptation is to find the correct derivative operator ∇ on $\mathbb{Z}$. In general, we make the following definition:

Definition 2.2. *A spatial derivative ∇ on $\mathbb{Z}$ is a linear operator in the space of functions on $\mathbb{Z}$ that maps any function f to another function ∇f, where, for each $k \in \mathbb{Z}$, there exists a coefficient $\alpha_k \in [0,1]$ such that*

$$(\nabla f)(k) = \alpha_k(\nabla^l f)(k) + (1-\alpha_k)(\nabla^r f)(k).$$

In other words, a derivative is defined by a family of coefficients ($\alpha_k \in [0,1]$), for $k \in \mathbb{Z}$. Each of these coefficients tells us how to mix, at a given point k, left and right derivatives. For example, the left (resp. right) derivative corresponds to the case where all the coefficients are equal to 1 (resp. 0).

First we show that a spatial derivative on $\mathbb{Z}$ for which there exists a fundamental solution to the n-translation problem must follow some necessary conditions. We next show that these necessary conditions allow us to reduce the translation problem to a more understandable problem of linear algebra in finite dimensions.

Proposition 2.3. *Fix integer $n \geq 1$ and a derivative ∇ on $\mathbb{Z}$ defined by a family of coefficients $(\alpha_k)_{k\in\mathbb{Z}}$. If there exists a solution μ_t to the n-translation problem (12), (13) associated with ∇ then $\alpha_0 = 0$ and $\alpha_n = 1$. Moreover, the support of μ_t is contained in $\{0,\dots n\}$.*

Proof. Let us first consider the function $f : \mathbb{Z} \to \mathbb{R}$ defined by $f(k) = 0$ if $k < 0$, and $f(k) = 1$ if $k \geq 0$. It is easy to show that $(\nabla f)(-1) = 1 - \alpha_{-1}$, $(\nabla f)(0) = \alpha_0$, and $(\nabla f)(k) = 0$ elsewhere. Let us now define the function $g : [0,1] \to \mathbb{R}$ by

$$g(t) := \sum_{k\in\mathbb{Z}} f(k)\mu_t(k) := \sum_{k\geq 0} \mu_t(k).$$

The initial and final conditions satisfied by μ_t show that $g(0) = 1 = g(1)$. On the other hand, the Equation (2) shows that

$$g'(t) = n\sum_{k\in\mathbb{Z}} \mu_t(k)\nabla f(k) = n[(1-\alpha_{-1})\mu_t(-1) + \alpha_0\mu_t(0)].$$

In particular $g'(t) \geq 0$. The fact that $g(0) = g(1)$ thus implies that $g'(t) = 0$ for every $t \in [0,1]$, and the condition $g'(0) = 0$ can be written $\alpha_0 = 0$. Moreover, the fact that $g(t) = 1$ for every $t \in [0,1]$ implies

$$\sum_{k\geq 0} \mu_t(k) = 1,$$

so μ_t is supported on $\mathbb{Z}_+$.

If we apply the same arguments to the function f defined by $f(k) = 1$ if $k \leq n$, and $f(k) = 0$ if $k > n$, we find that $\alpha_n = 1$, and that μ_t is supported on $\{k \in \mathbb{Z} \mid k \leq n\}$. □

An interesting consequence of Proposition 2.3 is that the translation problem of Equation (13) can be restricted to μ_t supported on $[0,n]$. That is, we can replace (13) by

$$\frac{\partial}{\partial t}\sum_{k=0}^{n} f(k)\mu_t(k) = n\sum_{k=0}^{n} \nabla f(k)\mu_t(k). \tag{14}$$

Now, let us consider the canonical basis $\mathcal{CB} := (e_0, \dots e_n)$ of the linear space of functions $\{0, \dots, n\} \to \mathbb{R}$. Let $X(t)$ be the column vector representing μ_t in $\mathcal{CB}$ (probability measures are canonically identified with functions), ie for every $k \in \{0, \dots n\}$, $(X(t))_k := \mu_t(k)$. The initial (resp. final) condition $\mu_0 = \delta_0$ (resp. $\mu_1 = \delta_n$) is equivalent to $X(0) = e_0$ (resp. $X(1) = e_n$). Moreover, Equation (14) is equivalent to the fact that for all vectors $Y \in M_{n,1}(\mathbb{R})$

$$\langle X'(t), Y\rangle = \frac{\partial}{\partial t}\langle X(t), Y\rangle = n\langle X(t), \nabla\rangle = n\langle \nabla^* X(t), Y\rangle, \tag{15}$$

where $\langle ., .\rangle$ is the usual (unweighted) scalar product on column vectors, and where ∇^* represents the adjoint with respect to this scalar product. This allows us to deduce that

$$X'(t) = n\nabla^* X(t), \tag{16}$$

and basic theorems on first-order linear differential systems thus allow us to write the n-translation problem:

Theorem 2.4. *Let $n \geq 1$ be an integer, and ∇ be a derivative on $\mathbb{Z}$, with $\alpha_0 = 0$ and $\alpha_n = 1$. Let A_∇ be the matrix of the linear operator ∇ in the canonical basis $\mathcal{CB}$. There exists a fundamental solution to the n-translation problem associated with ∇ if and only if, for every $t \in [0, 1]$, the column matrix*

$$X(t) := \exp(ntA_\nabla)e_0$$

has all its coefficients non-negative, and satisfies the final condition

$$X(1) = e_n. \tag{17}$$

The fundamental solution $\mu_t(k)$ is then given by $\mu_t(k) = (X(t))_k$.

We prove Theorem 1.2 using the properties of the spatial derivative ∇_n introduced in Definition 1.1. In this case we can be explicit about the form of ∇_n^*, and introduce a further map $\widetilde{\nabla}_n$ which will be used to prove Theorem 1.2 and the Poincaré inequality Theorem 1.3.

Definition 2.5.

1. *Let ∇_n^* be the adjoint operator of ∇_n for the unweighted scalar product on $l^2(\{0, \dots n\})$. We have the formula*

$$\nabla_n^* g(k) = \frac{1}{n}\left((n - k + 1)g(k - 1) - (n - 2k)g(k) - (k + 1)g(k + 1)\right),$$

where $g(-1) = g(n + 1) = 0$.

2. *We now fix $t \in (0, 1)$. Let $\widetilde{\nabla}_n$ be the adjoint operator of ∇_n for the scalar product with respect to the binomial law $b_{n,t}$ (taking $t \notin \{0, 1\}$ ensures that it is truly a scalar product on the space of functions $\{0, \dots n\} \to \mathbb{R}$)). We have:*

$$\begin{aligned}
\widetilde{\nabla}_n f(k) &= \frac{1}{b_{n,t}(k)}\nabla_n^*(f(k)b_{n,t}(k)) \\
&= \frac{n-k+1}{n}\frac{b_{n,t}(k-1)}{b_{n,t}(k)}f(k-1) - \frac{n-2k}{n}f(k) - \frac{k+1}{n}\frac{b_{n,t}(k+1)}{b_{n,t}(k)}f(k+1) \\
&= \frac{k}{n}\frac{1-t}{t}f(k-1) - \frac{n-2k}{n}f(k) - \frac{n-k}{n}\frac{t}{1-t}f(k+1).
\end{aligned} \tag{18}$$

The equivalence of the last two results follows since for all k,

$$\frac{b_{n,t}(k-1)}{b_{n,t}(k)} = \frac{k}{n-k+1}\frac{1-t}{t}.$$

We can relate properties of $\widetilde{\nabla}_n$ and ∇_n^* using conjugation by the linear operator D that maps any function $f : \{0, \ldots, n\} \to \mathbb{R}$ to the function Df defined by

$$\forall k \in \{0, \ldots, n\},\ Df(k) = b_{n,t}(k) f(k).$$

Moreover, as $t \in (0, 1)$, D is invertible and

$$\forall k \in \{0, \ldots, n\},\ D^{-1} f(k) = \frac{1}{b_{n,t}(k)} f(k).$$

This operator is useful to give a very simple relation between ∇_n^* and $\widetilde{\nabla}_n$:

$$\widetilde{\nabla}_n = D^{-1} \circ \nabla_n^* \circ D. \tag{19}$$

Proof of Theorem 1.2. We simply verify that (16) holds taking $X(t) = b_{n,t}(k)$ and ∇^* in the form given by Definition 2.5. We observe that in this case both sides of (16) have kth component equal to $b_{n,t}(k)\,(k/t - (n-k)/(1-t))$. The fact that $\frac{\partial}{\partial t} b_{n,t}(k)$ takes this form is immediate, and the corresponding result for the right hand side follows by Equations (18) and (19) since $n \frac{1}{b_{n,t}(k)} \nabla_n^* b_{n,t}(k) = n \widetilde{\nabla}_n \mathbf{1} = k/t - (n-k)/(1-t)$, where $\mathbf{1}$ denotes the function which is identically 1. □

3 Proof of the Poincaré inequality

From now on, we fix an integer $n \geq 1$, and we denote by ∇_n the finite difference operator of Definition 1.1. We recall the definition of the Krawtchouk polynomials from [11].

Theorem 3.1. *There exists a basis of polynomials in k, denoted $\phi_0, \ldots, \phi_n$, "laddered" (i.e. with $deg(\phi_r) = r$), and such that*

$$\sum_{k=0}^{n} \phi_r(k) \phi_s(k) b_{n,t}(k) = \frac{n! r!}{(n-r)!} \left(\frac{t}{1-t} \right)^r \delta_{rs} := C_{n,r} \delta_{rs}. \tag{20}$$

This family of polynomials is uniquely determined by the generating function in w

$$P(k, w) := \sum_{r=0}^{n} \frac{(1-t)^r}{r!} \phi_r(k) w^r = (1 + (1-t)w)^k (1 - tw)^{n-k}. \tag{21}$$

The discrete derivatives in k of $P(k, w)$ can be obtained by using the formulas

$$P(k-1, w) = P(k, w) \frac{1 - tw}{1 + (1-t)w} \text{ for all } k \geq 1 \tag{22}$$

$$P(k+1, w) = P(k, w) \frac{1 + (1-t)w}{1 - tw} \text{ for all } k \leq n-1 \tag{23}$$

Finally, since $\frac{\partial}{\partial w} w^r = r w^{r-1}$, we obtain

$$\sum_{r=0}^{n} \frac{(1-t)^r}{r!} r \phi_r(k) w^r = w \frac{\partial}{\partial w} P(k, w) = w P(k, w) \left(\frac{(1-t)k}{1 + (1-t)w} - \frac{t(n-k)}{1 - tw} \right). \tag{24}$$

Notice that ϕ_0 is the function identically equal to 1, and so $\nabla_n \phi_0 = 0$, which gives a sense to Proposition 3.2 when $r = 0$. To simplify the proof, we will define $\phi_{-1} = \phi_{n+1} = 0$.

Proposition 3.2. *For every $r \in \{0, \ldots, n\}$, we have*

1. *The operator ∇_n maps ϕ_r to a multiple of ϕ_{r-1}: $\nabla_n \phi_r = \dfrac{r(n-r+1)}{n(1-t)} \phi_{r-1}$.*

2. *The operator* $\widetilde{\nabla}_n$ *maps* ϕ_r *to a multiple of* ϕ_{r+1}*:* $\widetilde{\nabla}_n \phi_r = \dfrac{1}{nt}\phi_{r+1}$.

3. *The Krawtchouk polynomials are eigenfunctions for the linear map* $\left(\widetilde{\nabla}_n \circ \nabla_n\right)$*:*

$$\left(\widetilde{\nabla}_n \circ \nabla_n\right) \phi_r = \frac{r(n-r+1)}{n^2 t(1-t)} \phi_r.$$

Remark that these eigenvalues are not distinct, which does not allows us to deduce directly that the family $(\phi_0, \ldots, \phi_n)$ is a basis of the space of functions $\{0, \ldots, n\} \to \mathbb{R}$. This fact comes from the orthogonality with respect to the binomial scalar product.

Proof of Proposition 3.2. Part 1: It suffices to check the polynomial identity

$$\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \nabla_n \phi_r(k) w^r = \sum_{r=0}^{n} \frac{(1-t)^r}{r!} \frac{r(n-r+1)}{n(1-t)} \phi_{r-1}(k) w^r.$$

We will use the formula (21) to express both side of the last equation in terms of the polynomial $P(k, w)$. First, we have by Equations (22) and (23) that

$$\begin{aligned}
\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \nabla_n \phi_r(k) w^r &= \nabla_n P(k, w) \\
&= \frac{P(k, w)}{n} \left(k \left(1 - \frac{P(k-1, w)}{P(k, w)} \right) + (n-k) \left(\frac{P(k+1, w)}{P(k, w)} - 1 \right) \right) \\
&= \frac{P(k, w)}{n} \left(k \left(1 - \frac{1 - tw}{1 + (1-t)w} \right) + (n-k) \left(\frac{1 + (1-t)w}{1 - tw} - 1 \right) \right) \\
&= \frac{P(k, w)}{n} w \left(\frac{k}{1 + (1-t)w} + \frac{n-k}{1 - tw} \right).
\end{aligned}$$

For the right hand side, we have using (24) that

$$\begin{aligned}
\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \frac{r(n-r+1)}{n(1-t)} \phi_{r-1}(k) w^r &= \frac{w}{n} \sum_{r=0}^{n} \frac{(1-t)^r}{r!} (n-r) \phi_r(k) w^r \\
&= \frac{P(k, w)}{n} w \left(n - w \left(\frac{(1-t)k}{1 + (1-t)w} - \frac{t(n-k)}{1 - tw} \right) \right) \\
&= \frac{P(k, w)}{n} w \left(k \left(1 - \frac{(1-t)w}{1 + (1-t)w} \right) + (n-k) \left(1 + \frac{tw}{1 - tw} \right) \right) \\
&= \frac{P(k, w)}{n} w \left(\frac{k}{1 + (1-t)w} + \frac{n-k}{1 - tw} \right),
\end{aligned}$$

which gives the desired result.

Part 2: It suffices to check the polynomial identity

$$\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \widetilde{\nabla}_n \phi_r(k) w^r = \sum_{r=0}^{n} \frac{(1-t)^r}{r!} \frac{1}{nt} \phi_{r+1}(k) w^r.$$

Let us begin by studying the right hand side. Using the convention $\phi_{n+1} = 0$, we have by (24)

$$\begin{aligned}
\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \frac{1}{nt} \phi_{r+1}(k) w^r &= \frac{1}{nt(1-t)w} \sum_{r=0}^{n} \frac{(1-t)^{r+1}}{(r+1)!} (r+1) \phi_{r+1}(k) w^{r+1} \\
&= \frac{1}{nt(1-t)w} \sum_{r=0}^{n} \frac{(1-t)^r}{r!} r \phi_r(k) w^r \\
&= \frac{1}{nt(1-t)w} w \frac{\partial}{\partial w} P(k,w) \\
&= \frac{1}{nt(1-t)} P(k,w) \left(\frac{(1-t)k}{1+(1-t)w} - \frac{t(n-k)}{1-tw} \right).
\end{aligned}$$

The left hand side can be written

$$\sum_{r=0}^{n} \frac{(1-t)^r}{r!} \widetilde{\nabla}_n \phi_r(k) w^r = \widetilde{\nabla}_n P(k,w),$$

and we calculate using (22) and (23) that

$$\begin{aligned}
\widetilde{\nabla}_n P(k,w) &= P(k,w) \left(\frac{k}{n} \frac{1-t}{t} \frac{P(k-1,w)}{P(k,w)} - \frac{n-2k}{n} - \frac{n-k}{n} \frac{t}{1-t} \frac{P(k+1,w)}{P(k,w)} \right) \\
&= \frac{P(k,w)}{nt(1-t)} \left(k(1-t)^2 \frac{1-tw}{1+(1-t)w} - (n-2k)t(1-t) - (n-k)t^2 \frac{1+(1-t)w}{1-tw} \right) \\
&= \frac{P(k,w)}{nt(1-t)} \left((1-t)k \left(\frac{(1-t)(1-tw)}{1+(1-t)w} + t \right) - t(n-k) \left(\frac{t(1+(1-t)w)}{1-tw} + (1-t) \right) \right) \\
&= \frac{1}{nt(1-t)} P(k,w) \left(\frac{(1-t)k}{1+(1-t)w} - \frac{t(n-k)}{1-tw} \right),
\end{aligned}$$

and the proof is complete.

Part 3: follows directly by combining the two previous results. □

Similarly, there is another way to prove Part 2 of Proposition 3.2, using the properties of the exponential of the operator ∇_n^*:

Alternative proof of Proposition 3.2, Part 2.

$$\forall t \in [0,1], \ \exp(nt\nabla_n^*)(e_0) = (b_{n,t}(0), \ldots, b_{n,t}(n))^T. \tag{25}$$

The equation (21) allows us to show that the required result is equivalent to

$$\exp(nt(1-t)w\widetilde{\nabla}_n)(\phi_0) = (1+(1-t)w)^k (1-tw)^{n-k}. \tag{26}$$

As $\phi_0 = (1, \ldots, 1)^T$, the equation (25):

$$\begin{aligned}
D\phi_0 &= (b_{n,t}(0), \ldots, b_{n,t}(n))^T \\
&= \exp(nt\nabla_n^*)(e_0).
\end{aligned}$$

$$\begin{aligned}
\exp(nt(1-t)w\widetilde{\nabla}_n)(\phi_0) &= D^{-1} \exp(nt(1-t)w\nabla_n^*) D\phi_0 \\
&= D^{-1} \exp(nt(1-t)w\nabla_n^*) \exp(nt\nabla_n^*)(e_0) \\
&= D^{-1} \exp(nt(1+(1-t)w)\nabla_n^*)(e_0).
\end{aligned}$$

This means that, for every $k \in \{0, \ldots, n\}$:

$$\begin{aligned}
\exp(nt(1-t)w\widetilde{\nabla}_n)(\phi_0)(k) &= \frac{1}{b_{n,t}(k)} b_{n,t(1+(1-t)w)}(k) \\
&= \left(\frac{t(1+(1-t)w)}{t}\right)^k \left(\frac{1-t(1+(1-t)w)}{1-t}\right)^{n-k} \\
&= (1+(1-t)w)^k(1-tw)^{n-k}.
\end{aligned}$$

This proves the formula (26), and thus Part 2 of Proposition 3.2. □

We can complete the proof of Theorem 1.3, as follows:

Proof of Theorem 1.3. We can expand function $f(k) = \sum_{j=1}^n a_j \phi_j(k)$, since the assumption that $\sum_{k=0}^n f(k) b_{n,t}(k) = 0$ ensures that $a_0 = 0$. Using the normalization term $C_{n,r}$ from Equation (20), and the adjoint $\widetilde{\nabla}_n$ of Definition 2.5, we know that

$$\left(\widetilde{\nabla}_n \circ \nabla_n\right) f = \sum_{j=1}^n a_j \left(\widetilde{\nabla}_n \circ \nabla_n\right) \phi_j = \sum_{j=1}^n a_j \left(\frac{j(n-j+1)}{n^2 t(1-t)}\right) \phi_j, \tag{27}$$

by Part 3 of Proposition 3.2. This means that can write the RHS of Equation (3) as

$$\begin{aligned}
nt(1-t)\sum_{k=0}^n b_{n,t}(k) f(k) \left(\widetilde{\nabla}_n \circ \nabla_n\right) f(k) &= nt(1-t)\sum_{j=1}^n a_j^2 \frac{j(n-j+1)}{n^2 t(1-t)} C_{n,j} \\
&= \sum_{j=1}^n a_j^2 \frac{j(n-j+1)}{n} C_{n,j} \\
&\geq \sum_{j=1}^n a_j^2 C_{n,j}
\end{aligned}$$

which is the LHS of Equation (3). The inequality follows since $j(n-j+1)/n \geq 1$ with equality if and only if $j = 1$ or $j = n$. □

Remark 3.3. *Theorem 1.3 is a stronger result than Klaassen's Equation (5). This follows by noting that for any k, we can express*

$$\begin{aligned}
&(\nabla_n f)(k)^2 \\
&= \frac{n-k}{n}(f(k+1)-f(k))^2 + \frac{k}{n}(f(k)-f(k-1))^2 - \frac{k(n-k)}{n^2}(f(k+1)-2f(k)+f(k-1))^2 \\
&\leq \frac{n-k}{n}(f(k+1)-f(k))^2 + \frac{k}{n}(f(k)-f(k-1))^2.
\end{aligned}$$

This means that, using the fact that $\frac{k}{n} b_{n,t}(k) = \frac{n-k+1}{n}\frac{t}{1-t} b_{n,t}(k-1)$, we can express the RHS of (3) as

$$\begin{aligned}
&nt(1-t)\sum_{k=0}^n b_{n,t}(k)(\nabla_n f)(k)^2 \\
&\leq nt(1-t)\sum_{k=0}^n b_{n,t}(k)\left(\frac{n-k}{n}(f(k+1)-f(k))^2 + \frac{k}{n}(f(k)-f(k-1))^2\right) \\
&= t\sum_{k=0}^n b_{n,t}(k)(n-k)(f(k+1)-f(k))^2\left((1-t)+t\right),
\end{aligned}$$

which is the RHS of (5).

Remark 3.4. *The explicit form taken by the operators ∇_n and $\widetilde{\nabla}_n$ and their action on the Krawtchouk polynomials suggest possible links with the Ehrenfest urn model, see for example [8]. It is possible to recover the Poincaré inequality Theorem 1.3 from results about this well-studied model, as follows. For $n \geq 1$ and $0 < t < 1$ we consider the Markov chain on $\{0, \dots n\}$ generated by the operator L defined by*

$$Lf(k) := k(1-t)[f(k-1) - f(k)] + (n-k)t[f(k+1) - f(k)]. \tag{28}$$

The invariant measure for this reversible Markov chain is the binomial measure $b_{n,t}$. Moreover, the spectrum of the operator L has been completely described by Karlin and McGregor in [8] who show that, for $r = 0, \dots n$ we have $L\phi_r = -r\phi_r$. Further, $\widetilde{\nabla}_n \circ \nabla$ is related to L by the simple formula

$$t(1-t)\widetilde{\nabla}_n \circ \nabla_n = -\frac{1}{n^2}L^2 - \frac{n+1}{n^2}L, \tag{29}$$

which can be verified directly. Applying equation (29) to each eigenfunction ϕ_r gives

$$(1-t)t\widetilde{\nabla}_n \circ \nabla_n \phi_r = \frac{-r^2 + (n+1)r}{n^2}\phi_r = \frac{r(n-r+1)}{n^2}\phi_r, \tag{30}$$

which allows us to recover (27), which implies the Poincaré inequality Theorem 1.3. This proof seems simpler and more direct; however, it does not give us information about the action of the operators ∇_n and $\widetilde{\nabla}_n$ on the family of Krawtchouk polynomials.

We can understand the decomposition discussed in Remark 3.3 in this context, by noting that (29) means we can write

$$nt(1-t)\sum_{k=0}^{n} b_{n,t}(k) \left(\nabla_n f\right)(k)^2 = -\sum_{k=0}^{n} b_{n,t}(k) Lf(x) f(k) - \frac{1}{n}\sum_{k=0}^{n} b_{n,t}(k) \left(Lf(k) + f(k)\right) Lf(k),$$

where further calculations show that the first term is equal to Klaasen's bound, given as the RHS of (5).

References

[1] S. Bobkov and F. Götze. Discrete isoperimetric and Poincaré-type inequalities. *Probability Theory and Related Fields*, 114:245–277, 1999.

[2] S. G. Bobkov. Some extremal properties of the Bernoulli distribution. *Teor. Veroyatnost. i Primenen.*, 41(4):877–884, 1996.

[3] S. G. Bobkov. An isoperimetric inequality on the discrete cube, and an elementary proof of the isoperimetric inequality in Gauss space. *Ann. Probab.*, 25:206–214, 1997.

[4] S. G. Bobkov and M. Ledoux. On modified logarithmic Sobolev inequalities for Bernoulli and Poisson measures. *J. Funct. Anal.*, 156(2):347–365, 1998.

[5] T. Cacoullos. On upper and lower bounds for the variance of a function of a random variable. *Ann. Probab.*, 10(3):799–809, 1982.

[6] L. H. Y. Chen and J. H. Lou. Characterization of probability distributions by Poincaré-type inequalities. *Ann. Inst. H. Poincaré Probab. Statist.*, 23(1):91–110, 1987.

[7] H. Chernoff. A note on an inequality involving the normal distribution. *Ann. Probab.*, 9(3):533–535, 1981.

[8] S. Karlin and J. McGregor. Ehrenfest urn models. *Journal of Applied Probability*, 2(2):352–376, 1965.

[9] C. Klaassen. On an inequality of Chernoff. *Ann. Probab.*, 13(3):966–974, 1985.

[10] L. Saloff-Coste. Lectures on finite Markov Chains. In P. Bernard, editor, *Lectures on Probability Theory and Statistics, St-Flour 1996*, number 1665 in Lecture Notes in Mathematics, pages 301–413. Springer-Verlag, 1997.

[11] G. Szegő. *Orthogonal Polynomials.* American Mathematical Society, New York, revised edition, 1958.