6 Canonical representations of factorization theorems for special sums

6.1 Rationale and motivating the discussion

There is a natural question that works its way into the analysis of the prior research and publications by Merca and Schmidt that we have summarized above. One reflection, in hindsight, as to why these seemingly simple expansions related to Lambert series generating functions resulted in so many acceptances in excellent journals is that they make special, and as at least one reviewer had pointed out, rare connections between classically multiplicative-only constructions and the theory of partitions. Prior to those several publications, only Andrews and a handful of other authors had found such relations, and none yet it seems so general and clear cut. The choice of factorizing the Lambert series OGF expansions by inserting a multiple of the generating function for the partition function, $p(n) = [q^n](q;q)_\infty^{-1}$, leads to a representation for the matrices $s_{n,k}$ and $s_{n,k}^{-1}$ that both involve special functions that are central to the theory of partitions [?]. The initial relation of the $s_{n,k}$ to partition theoretic constructions was proved independently by Merca in 2017 [?], and seen from a different lense in [19] near the same time we decided to collaborate on generalizing this material.

The divisor sums, $f * 1$, that are enumerated as coefficients of the LGF of $f(n)$ are only one of a larger class of special sums that we have seen, and decided to explore using these factorization type theorems for their OGFs. The original LGF case suggests that there should (á priori) be a more natural, or even a "best possible" (in some senses) canonical expression for how we choose the reciprocal generating function factor of the functions $\mathcal{C}(q)$. Indeed, as we showed in [?], one can seemingly always draw connections to the standard (Euler) partition function by choosing $\mathcal{C}(q) := (q;q)_\infty$, but the representation for the ordinary (non-inverse) matrices is decidely less expressive, telling nor as natural with respect to what new properties this lets us view in the struture of these summation based discrete convolution sequences. Hence, we seek a better way to write these invertible, lower triangular matrix based OGF expressions so that we have the most natural interpretation of the resulting coefficient identities for the particular application at hand. The resulting expressions should be considered optimal, or "best possible", in a way we can use to non-trivially quantify the most statistically relevant choice of $\mathcal{C}(q)$. There is a formal, and less qualitatively constructive starting point, for actually grasping a tenuous description of the relevance behind our new OGF based interpretations that formally enumerate the class of special convolution type sums defined in (45) below.

The material we present in this concluding section of the thesis is not exhaustive, nor conclusive. Rather it serves to motivate a discussion of rigorously formulating "best possible" factorization theorems and relationships between application-dependent convolution type sequences. These so-termed "canonical" expressions that arise in other important applications and future useful constructions based on our work from Section 3. We will give visual intuition for why that relationship was special for the LGF series examples and seek to formalize a rigorous metric, or meaningful correlation statistic, that can be used to generalize the qualitative "niceness" we found in the LGF factorization theorem sequences. A few conjectures are presented in the last subsection below that suggest a loose application-tied partition theoretic interpretation behind the ideal correlation for factorization theorems we have for a class of more general convolution type sums (defined precisely in the next subsection).

6.2 Definitions and constructions of generalized kernel-based convolution type sums

For a bivariate kernel function $\mathcal{D} : \mathbb{N}^2 \to \mathbb{C}$, we define the next class of convolution type sums ($\mathcal{D}$-convolution sums) according to the formula

$$(f \boxdot_{\mathcal{D}} g)(n) := \sum_{k=1}^{n} f(k)g(n+1-k)\mathcal{D}(n,k), n \geq 1. \tag{45}$$

In the section ahead, we are able to connect these special convolution type sums that form a widely reaching class of applications through particular specializations of the $(f, g; \mathcal{D})$.

Definition 6.1. We say that a kernel, or weight function $\mathcal{D} : (\mathbb{Z}^+)^2 \to \mathbb{C}$ is *lower triangular* if $\mathcal{D}(n,k) = 0$ for all $k > n \geq 1$. We say that this kernel function is invertible provided that

$$\det\left[(\mathcal{D}(n,k))_{1 \leq n,k \leq N}\right] \neq 0, \forall N \geq 1.$$

Suppose that $\mathcal{D}$ is an invertible, lower triangular kernel function, and that the arithmetic function g is invertible with respect to $\mathcal{D}$-convolution, i.e., defined such that $g(1) \neq 0$. Then we can express the OGF of these sums according to the following parameterized invertible matrix based factorizations for $n \geq 1$:

$$(f \boxdot_{\mathcal{D}} g)(n) = [q^n]\frac{1}{\mathcal{C}(q)} \times \sum_{n\geq 1}\left(\sum_{k=1}^{n} s_{n,k}(g; \mathcal{C}, \mathcal{D}) f(k)\right) q^n, \text{ for } \mathcal{C}(0) \neq 0. \tag{46}$$

Proposition 6.2 (Inversion)**.** *Suppose that $\mathcal{D}$ is a kernel function for a convolution sequence defined in* (45) *that is both invertible and satisfies $\mathcal{D}(n,n) \neq 0$ for all $n \geq 1$. Let the corresponding lower triangular sequence of entries for its inverse matrix be denoted by $\mathcal{D}^{-1}(n,k)$. Then for any $n \geq 1$ we have that*

$$g(n) = (f \boxdot_{\mathcal{D}} 1)(n) \iff f(n) = \sum_{k=1}^{n} g(k)\mathcal{D}^{-1}(n,k).$$

Proof. By construction, we suppose that $1 \leq N < +\infty$ and that we have two N-dimensional vectors, $\vec{f} := [f(1), \ldots, f(N)]^T$ and $\vec{g} := [g(1), \ldots, g(N)]^T$. It follows that

$$\vec{g} = \mathcal{D}(N) \cdot \vec{f} \implies \vec{f} = \mathcal{D}(N)^{-1} \cdot \vec{g},$$

where $\mathcal{D}(N) := (\mathcal{D}(n,k))_{1\leq n,k\leq N}$. $\square$

Suppose that $\mathcal{D}(n,k)$ is an invertible, lower triangular kernel function. We say that an arithmetic function g is invertible with respect to $\mathcal{D}$-convolution if there exists a (left) inverse function $g^{-1}[\mathcal{D}](n)$ such that for all integers $n \geq 1$ we have that $(g^{-1}[\mathcal{D}] \boxdot_{\mathcal{D}} g)(n) = \delta_{n,1}$. We can restrict ourselves to the cases where we take $\mathcal{D}$ to be *symmetric* with respect to $\mathcal{D}$-convolution: That is, where we have that $\mathcal{D}(n,k) = \mathcal{D}(n, n+1-k)$ for all $1 \leq k \leq n$. In these cases, we have that the left and corresponding right inverse functions of any g with respect to $\mathcal{D}$-convolution are identical when they exist.

Proposition 6.3 (Inverses of an arithmetic function with respect to $\mathcal{D}$-convolution)**.** *An arithmetic function g is invertible with respect to $\mathcal{D}$-convolution for a fixed invertibly lower triangular and symmetric kernel $\mathcal{D}$ if and only if $g(1) \neq 0$. When the function $g^{-1}[\mathcal{D}](n)$ exists, it is unique, and can be computed exactly by recursion via the following formula:*

$$g^{-1}[\mathcal{D}](n) = \begin{cases} \frac{1}{\mathcal{D}(1,1)g(1)}, & n = 1; \\ -\frac{1}{\mathcal{D}(n,n)g(1)} \times \sum\limits_{1\leq k<n} g^{-1}[\mathcal{D}](k) g(n+1-k) \mathcal{D}(n,k), & n \geq 2. \end{cases}$$

Moreover, provided that $g^{-1}[\mathcal{D}]$ exists, we have that

$$g^{-1}[\mathcal{D}](n) = \mathcal{D}^{-1}(n,1) \times [q^{n-1}]\left(\sum_{n\geq 0} g(n+1) q^n\right)^{-1}, n \geq 1.$$

Proof. Fix any arithmetic function g with $g(1) \neq 0$. The recursive formula follows by a rearrangement of the terms in the equation $(g^{-1}[\mathcal{D}] \boxdot_{\mathcal{D}} g)(n) = \delta_{n,1}$. To prove the exact formula, we see that we can set up an invertible matrix vector system of the form $A_{\mathcal{D},g}(N) \cdot \vec{f} = \vec{b}$ and solve for $\vec{f}$ where $A_{\mathcal{D},g}(N) := (\mathcal{D}(n,k) g(n+1-k) \, [k \leq n]_{\delta})_{1\leq n,k\leq N}$, $\vec{f} = [g^{-1}[\mathcal{D}](1), \ldots, g^{-1}[\mathcal{D}](N)]^T$, and $\vec{b} = [1, 0, \ldots, 0]^T$ for any $N \geq 1$. Then we have that

$$g^{-1}[\mathcal{D}](n) = (A^{-1}_{\mathcal{D},g}(N))_{n,1}, \forall N \geq n \geq 1.$$

Notice that for $\mathcal{D}(N) := (\mathcal{D}(n,k))_{1\leq n,k\leq N}$, we can write

$$A_{\mathcal{D},g}(N) = \mathcal{D}(N) \cdot \begin{bmatrix} g(1) & 0 & 0 & \cdots & 0 \\ g(2) & g(1) & 0 & \cdots & 0 \\ g(3) & g(2) & g(1) & \cdots & 0 \\ & & & \vdots & 0 \\ g(N) & g(N-1) & g(N-2) & \cdots & g(1) \end{bmatrix},$$

where the right-hand-side matrix involving g is an invertible Topelitz matrix. Thus by inversion, we see that our claimed formula is correct. $\square$

A table of inverse functions $g^{-1}[\mathcal{D}]$ for any fixed arithmetic function g such that $g(1) \neq 0$, and any fixed invertible, lower triangular kernel function satisfying $\mathcal{D}(n,n) \neq 0$ for all $n \geq 1$, is computed symbolically in the listings below.

n	$g^{-1}[\mathcal{D}](n)$
1	$\frac{1}{\mathcal{D}(1,1)g(1)}$
2	$-\frac{\mathcal{D}(2,2)g(2)}{\mathcal{D}(1,1)\mathcal{D}(2,1)g(1)^2}$
3	$\frac{\mathcal{D}(2,2)\mathcal{D}(3,2)g(2)^2}{\mathcal{D}(1,1)\mathcal{D}(2,1)\mathcal{D}(3,1)g(1)^3} - \frac{\mathcal{D}(3,3)g(3)}{\mathcal{D}(1,1)\mathcal{D}(3,1)g(1)^2}$
4	$-\frac{\mathcal{D}(2,2)\mathcal{D}(3,2)\mathcal{D}(4,2)g(2)^3}{\mathcal{D}(1,1)\mathcal{D}(2,1)\mathcal{D}(3,1)\mathcal{D}(4,1)g(1)^4} + \frac{(\mathcal{D}(2,1)\mathcal{D}(3,3)\mathcal{D}(4,2)+\mathcal{D}(2,2)\mathcal{D}(3,1)\mathcal{D}(4,3))g(3)g(2)}{\mathcal{D}(1,1)\mathcal{D}(2,1)\mathcal{D}(3,1)\mathcal{D}(4,1)g(1)^3} - \frac{\mathcal{D}(4,4)g(4)}{\mathcal{D}(1,1)\mathcal{D}(4,1)g(1)^2}$

In what follows, we adopt the notation that $c_n(\mathcal{C}) := [q^n]\mathcal{C}(q)$ and $p_n(\mathcal{C}) := [q^n]\mathcal{C}(q)^{-1}$ for any $n \geq 0$ and any OGF $\mathcal{C}(q)$ such that $\mathcal{C}(0) \neq 0$.

Theorem 6.4 (Generalized factorization theorems for $\mathcal{D}$-convolution)**.** *Suppose that g is any arithmetic function that is invertible with respect to $\mathcal{D}$-convolution for some fixed invertible, lower triangular kernel function $\mathcal{D}(n,k)$. Then the matrices with entries given by $s_{n,k}(g;\mathcal{C},\mathcal{D})$ in (46) are invertible and satisfy the following formulas for $1 \leq k \leq n$:*

$$s_{n,k}(g;\mathcal{C},\mathcal{D}) = \sum_{j=1}^{n} c_{n-j}(\mathcal{C})g(j+1-k)\mathcal{D}(j,k)$$
$$s_{n,k}^{-1}(g;\mathcal{C},\mathcal{D}) = \sum_{j=1}^{n} g^{-1}[\mathcal{D}](n+1-j)p_{j-k}(\mathcal{C})\mathcal{D}(n,j).$$

Proof. The formula for the ordinary matrix entries is obvious upon multiplying both sides of (46) by the OGF $\mathcal{C}(q)$, and then taking the coefficients of q^n and $f(k)$ in the resulting expansion. The proof of the inverse matrix formulas is routine, but less obvious. Since $s_{n,k}(g;\mathcal{C},\mathcal{D})$ is lower triangular with non-zero entries when $n = k$ for all $n \geq 1$, it forms a sequence of invertible square matrices taking determinants over $1 \leq n, k \leq N$ for each fixed $N \geq 1$. Consider the special case of the $\mathcal{D}$-convolution sums where $f(n) := s_{n,k}^{-1}(g;\mathcal{C},\mathcal{D})$, which we know exists uniquely, for each fixed $k \geq 1$. By the orthogonality relations between the lower triangular ordinary and inverse matrices, we can see that

$$p_{n-k}(\mathcal{C}) = \sum_{j=1}^{n} s_{n,k}^{-1}(g;\mathcal{C},\mathcal{D})g(n+1-j)\mathcal{D}(n,j).$$

Since g is invertible with respect to $\mathcal{D}$-convolution, we recover our claimed formula for the inverse matrix entries. $\square$

Question 6.5. The question of which OGF $\mathcal{C}(q)$ we should choose in expanding the representations of the generating functions for the $\mathcal{D}$-convolution sums in (46) boils down to formally quantifying which coefficients of this OGF are the most qualitatively expressive with respect to a fixed kernel function $\mathcal{D}(n,k)$. We elaborate on this question posed by Professor Michael Lacey that remains from my August 2020 oral exam presentation at Georgia Tech in the next subsections below.

6.3 Rigorous statistics for correlation-based quantification of the "canonically best" property

There is a vast body of modern literature in number theory that motivates semi-standardized ways to quantify relationships between functions and sequences we study via correlation based statistics. There is historically relevant literature about using statistical analysis to motivate studying number theoretic objects. For example, the non-trivial zeros of the Riemann zeta function have been related and bounded via pair correlation formulas. Moreover, this topic continues to be a active and fruitful way of understanding this complicated subject matter. We recall from [**?**] that results in analytic number theory that make sense of the distribution of the non-trivial zeros of $\zeta(s)$ originated in the work of Montgomery. Subsequent follow-up work that collectively builds on Montgomery's contributions in the context of L-functions, Gaussian Unitary Ensemble (or GUE), applications in random matrix theory and their associated correlation statistics is famously due to Hejal, Rudnick, Sarnak and Odlyzko.

We posit by extension that using correlation metrics, or so-called sequence correlation *statistics*, to precisely define and rigorously formulate what we consider to be best possible attainable relationships for the factorization theorems given in (46). In general, we can study the so-called *sequence vector correlation* (including the information theoretic cross-correlation statistics seen below) that relate more general sequences and vectors of real and rational numbers. In our case, we need to identify and prove optimal representations for our notion of the "best possible", or optimal. e.g., canonical view point, for how we should express the OGF factorization theorems as they are identified in (46). We then set out to precisely construct formulas that can be maximized (minimized) with respect to all possible one-dimensional sequences in a way that captures the qualitatively meaningful relationships between the sequences from the LGF case. The goal is to do this in a very general setting that reveals underlying hidden relationships charaterizing any particular class of $\mathcal{D}$-convolution type sums in analog to the observations of natural relationships between multiplicative number theory and the partition functions from the LGF case witnessed in Section 3.

6.3.1 Defining precise correlation statistics

Example 6.6 (A model starting point)**.** The exact bounded ranges we can expect for cross-correlation coefficients to express a numerical index between vectors in our problem context are, in general, variable and subject to the qualitative interpretation which we have to reason about separately to ensure a good model fit. If we wish to normalize the range to be within $[-1,1]$, there is the standardized definition of the (non-central, or non-centralized) *Pearson correlation coefficient.* It is defined as the numerical statistic relating any two N-tuples, $\vec{a} := (a_1,\ldots,a_N), \vec{b} := (b_1,\ldots,b_N) \in \mathbb{Q}^N$ for any fixed $N \geq 1$, given by

$$\operatorname{PearsonCorr}(N;\vec{a},\vec{b}) := \frac{1}{N} \times \frac{\sum\limits_{j=1}^{N} a_j b_j}{\sqrt{\sum\limits_{1\leq i,j\leq N} a_i^2 b_j^2}} \in [-1,1].$$

Notation 6.7. We again define the shorthand sequence notation of $c_n(\mathcal{C}) := [q^n]\mathcal{C}(q)$ and $p_n(\mathcal{C}) := [q^n]\mathcal{C}(q)^{-1}$ for any $\mathcal{C}(q)$ such that $\mathcal{C}(0) \neq 0$. We are going to adapt the non-centralized Pearson cross-correlation formula by choosing our correlation statistic to be computed according to the following sums:

$$\begin{aligned} \operatorname{Corr}(n;\mathcal{C},\mathcal{D}) &:= \frac{1}{n} \times \frac{\sum\limits_{k=1}^{n} |c_k(\mathcal{C})\mathcal{D}^{-1}(n,k)|}{\sqrt{\left(\sum\limits_{k=1}^{n} c_k(\mathcal{C})^2\right)\left(\sum\limits_{k=1}^{n} \mathcal{D}^{-1}(n,k)^2\right)}} \\ \operatorname{Corr}(\mathcal{C},\mathcal{D}) &:= \sum_{n\geq 1} \operatorname{Corr}(n;\mathcal{C},\mathcal{D}). \end{aligned} \tag{47}$$

Question 6.8 (The crux of our correlation statistic optimization problem)**.** For a fixed lower triangular, invertible kernel function $\mathcal{D}$, we need to identify a concrete candidate OGF, $\mathcal{C}(q)$, so that

$$0 \leq \operatorname{Corr}(\mathcal{C},\mathcal{D}) < +\infty,$$

is maximized or minimized (and finite) over all possible input functions $\mathcal{C}(q) \in \mathbb{Q}[[q]]$ such that $\mathcal{C}(0) \neq 0$, or alternately $\mathcal{C}(q) \in \mathbb{Z}[[q]]$ with $\mathcal{C}(0) = 1$. Note that this criteria and the corresponding maximization procedure is always independent of the arithmetic functions f, g input to the weighted $\mathcal{D}$-convolution sums, $f \boxdot_{\mathcal{D}} g$.

Example 6.9 (Finding optimal statistics for the LGF case)**.** We will make a somewhat arbitrary decision that works well in practice to define

$$f(n,k) := \begin{cases} \frac{1}{n} \times \frac{c_k(\mathcal{C})\mathcal{D}^{-1}(n,k)}{\left(\sum\limits_{m \leq n} c_m(\mathcal{C})^2\right)^{\frac{1}{2}} \left(\sum\limits_{m \leq n} \mathcal{D}^{-1}(n,m)^2\right)^{\frac{1}{2}}}, & \text{if } 1 \leq k \leq n \leq N; \\ 0, & \text{otherwise,} \end{cases}$$

Notice that in the cases we next look at for the LGF example, we have that

$$\sum_{m \leq n} \mathcal{D}^{-1}(n,m)^2 = \sum_{d|n} \mu^2(d) = 2^{\omega(n)}, n \geq 1.$$

We then want to optimize the minimal bounded formulas

$$\lim_{N \to \infty} \sum_{n \leq N} \sum_{k=1}^{n} f(n,k) = \lim_{N \to \infty} \sum_{n \leq N} \frac{(|c_{-}(\mathcal{C})| * |\mu|)(n)}{n\rho_{\mathcal{C}}(n) \left(\sqrt{2}\right)^{\omega(n)}} \in [0,1],$$

over all $\mathcal{C}(q)$ such that $\mathcal{C}(0) \neq 0$ and with

$$\rho_{\mathcal{C}}(n) := \left(\sum_{0 \leq m \leq n} c_m(\mathcal{C})^2\right)^{\frac{1}{2}}.$$

We see that minimizing the reciprocal of the limiting series in the previous equation leads to a maximal possible bound on the cross-correlation statistics we defined in (47). The preliminary numerical results we cite for this case in Section 6.4 below is able to numerically predict how closely this statistic for the LGF case comes to attaining the theoretically maximal correlation statistic in limiting cases. These series approximations can be made very accurate for certain classes of OGFs that commonly arise in applications.

6.3.2 Example: Visual projections of intuition for the LGF sequence cases

We will project this sense of distortion (versus similarity) between two tuples of values (truncated vectors of OGF coefficients) onto an easy to spot image. This visualization relies on how clearly the projected data allows us to look at *Tux*, the classic good-luck-forebearing Linux penguin mascot, depicted below in his traditional Linux kernel emblem modified here to assist with distinguishing features in the correlation statistic values. We use a variant of the built-in image processing functions within modern releases of *Mathematica* to generate these images [**?**]. The representative source code is reproduced from the thesis software package repository on `GitHub` at the link listed in the citations. The results displayed in Figure 6.1 serve to illustrate the point (shown on page 63). Amongst those selected $\mathcal{C}(q)$ that satisfy $\mathcal{C}(0) \notin \{0, \pm\infty\}$ and that have integer-valued coefficients, we conclude by visual inspection that our choice of $\mathcal{C}(q) := (q;q)_\infty$ appears optimal!

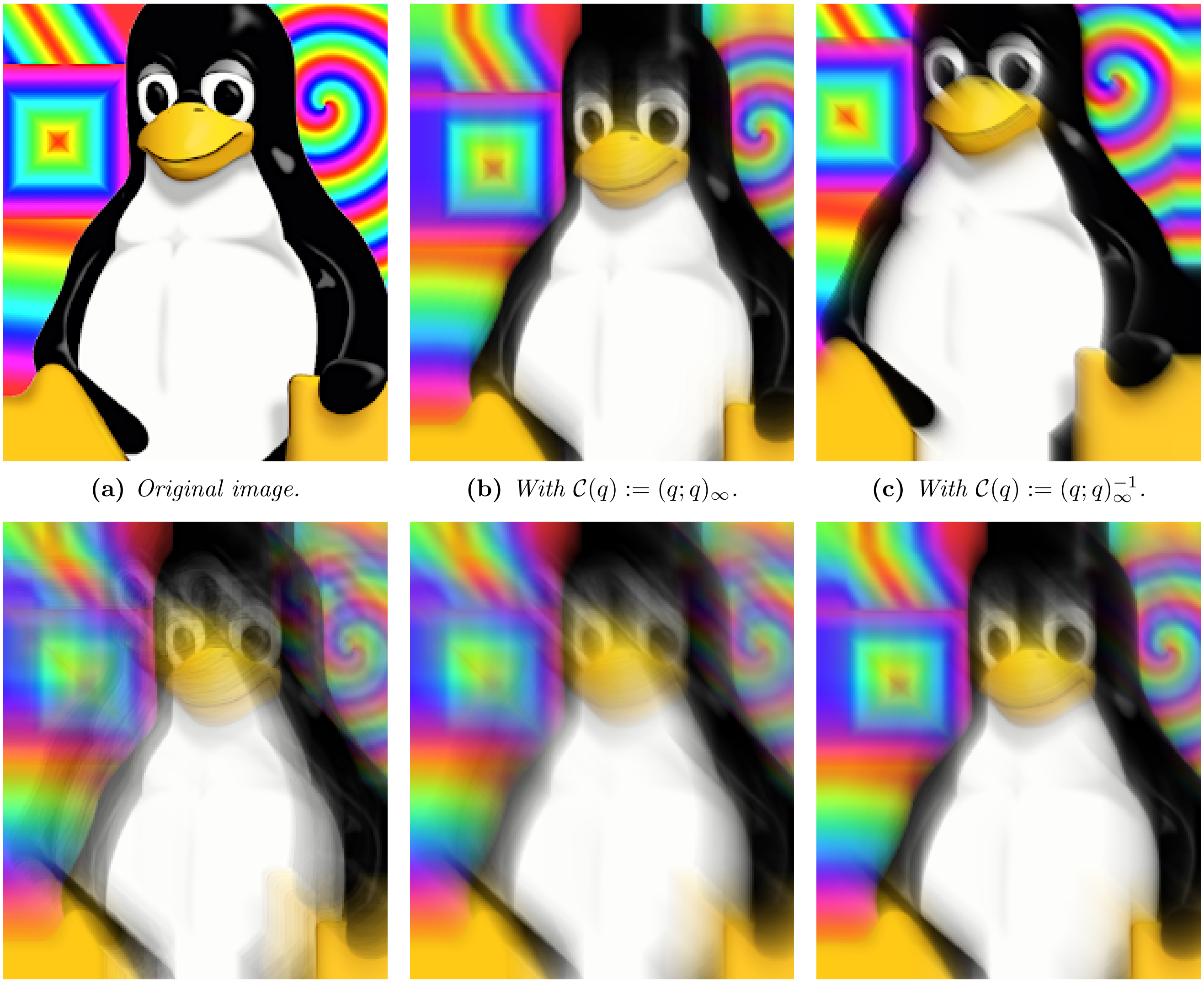
6.4 Maximal correlation bounds for the LGF case

Since the task of identifying the target limiting cross-correlation statistic in (47) is substantially complicated in the general case, we first look at the problem of optimality for the LGF case. For each such OGF $\mathcal{C}(q)$, we define

$$\operatorname{Corr}_{\mathrm{LGF}}(\mathcal{C}) = \lim_{n \to \infty} \sum_{k=1}^{n} \frac{\mu^2(k)}{k} \times \sum_{j \leq \left\lfloor \frac{n}{k} \right\rfloor} \frac{|c_j(\mathcal{C})|}{j\rho_{\mathcal{C}}(jk)(\sqrt{2})^{\omega(jk)}} = \sum_{j,k \geq 1} \frac{\mu^2(j)|c_k(\mathcal{C})|}{(jk)\rho_{\mathcal{C}}(jk)(\sqrt{2})^{\omega(jk)}}, \tag{48}$$

where we define the partial variance of $\mathcal{C}$ to be

$$\rho_{\mathcal{C}}(N) := \sqrt{\sum_{1 \leq i \leq N} c_i(\mathcal{C})^2}, N \geq 1.$$



(a) *Original image.* **(b)** *With* $\mathcal{C}(q) := (q;q)_\infty$. **(c)** *With* $\mathcal{C}(q) := (q;q)_\infty^{-1}$.

(d) *With* $\mathcal{C}(q) := (q^2;q^5)_\infty$. **(e)** *With* $\mathcal{C}(q) := (1-q)^{-3/2}$. **(f)** *With* $\mathcal{C}(q) := (1-q)^{-1}$.

Figure 6.1: *Correlation statistics projected onto the image of Tux for various choices of the OGF* $\mathcal{C}(q)$ *(with integer coefficients) to show a visual comparison of how well related the corresponding sequences are. Distortions of the original image indicate a less well correlated set of sequences corresponding to that choice of the* $\mathcal{C}(q)$ *that defines the precise form of the LGF factorization theorem coefficients.*

The previous doubly infinite series for the LGF correlation statistic is non-trivial to tightly bound from above and below because we have, in general, for $j, k \geq 1$ that

$$\omega(jk) = \omega(j) + \omega\left(\frac{k}{(k,j)}\right).$$

That is, the function $\omega(n)$ is only *strongly* (as opposed to *completely*) *additive.*

Definition 6.10. Fix any $0 < \delta < +\infty$. Provided an input test OGF $\mathcal{C}(q)$, we set

$$A_0(\mathcal{C}, \delta) := \lim_{N\to\infty} \frac{1}{N^\delta} \times \sqrt{\sum_{1\leq n\leq N} c_n(\mathcal{C})^2}.$$

We are naturally interested in finding the optimal, or so-termed "canonically best" correlation coefficient that corresponds to an explicit OGF $\mathcal{C}(q)$. We have already noticed that taking $\mathcal{C}(q) := (q;q)_\infty$ leads to very interesting relationships between the matrices in the Lambert series factorization theorems. Conjecture 6.17 given at the end of this section is suggestive of why the apparently optimal OGF $\mathcal{C}(q)$ has series coefficients that satisfy (by the *pentagonal number theorem*)

$$\delta = \sup\left\{\rho > 0 : A_0(\mathcal{C}, \rho) > 0\right\} \equiv \frac{1}{4},$$

in the last definition. Hence, we are interested in considering OGFs $\mathcal{C}(q)$ with integer coefficients such that the maximal $\delta > 0$ in the definition for which $A_0(\mathcal{C}, \delta) > 0$ is given by $\delta := \frac{1}{4}$.

Using the same construction as the special case where $\mathcal{C}(q) := (q;q)_\infty$, OGF forms whose coefficients are in $\{0, \pm 1\}$ such that the corresponding $\delta = \frac{1}{4}$ can be seen as often having non-zero coefficients, $|c_n(\mathcal{C})|$, for $n \in \{p(n)\}_{n=-\infty}^{\infty} \backslash \{0\}$ where $\hat{p}(n) = \frac{n(an+b)}{2}$ for integers $a \geq 3$ and $1 \leq b < a$. As we can see through the next OGF examples, generating functions of this form are natural to consider in partition theoretic applications [8, §19.9]:

$$(q;q)_\infty = \prod_{n\geq 0}\left\{(1-q^{3n+1})(1-q^{3n+2})(1-q^{3n+3})\right\} = \sum_{n=-\infty}^{\infty} (-1)^n q^{n(3n+1)/2} \quad \text{(A)}$$

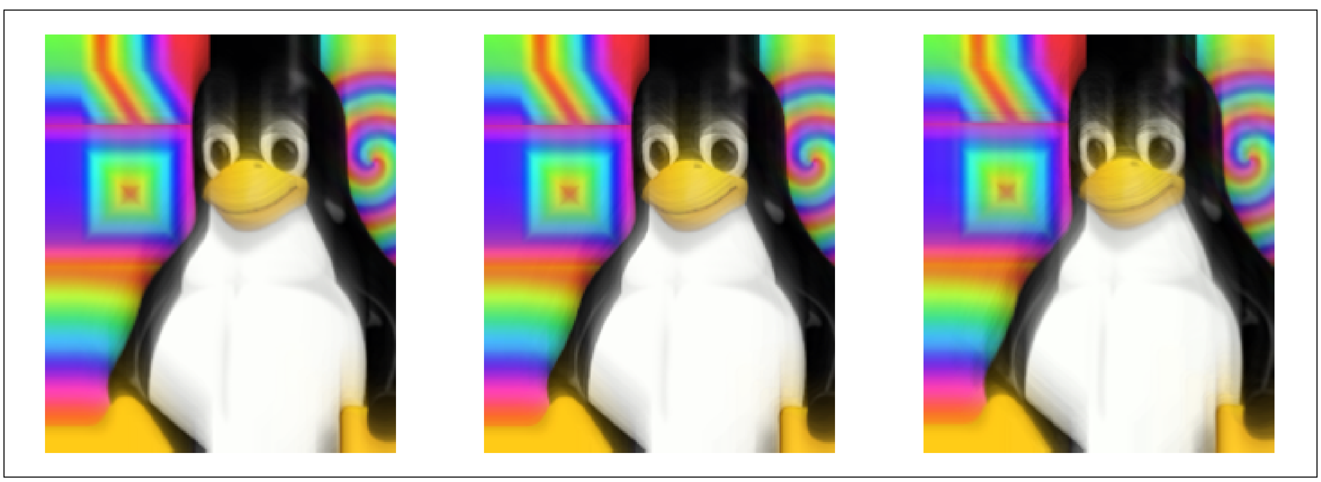
$$(-q;q^2)_\infty^2 (q^2;q^2)_\infty = \prod_{n\geq 0}\left\{(1+q^{2n+1})^2(1-q^{2n+2})\right\} = \sum_{n=-\infty}^{\infty} q^{n^2}$$

$$(q;q^2)_\infty^2 (q^2;q^2)_\infty = \prod_{n\geq 0}\left\{(1-q^{2n+1})^2(1-q^{2n+2})\right\} = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2}$$

$$(q;q^5)_\infty (q^4;q^5)_\infty (q^5;q^5)_\infty = \prod_{n\geq 0}\left\{(1+q^{5n+1})(1-q^{5n+4})(1-q^{5n+5})\right\} = \sum_{n=-\infty}^{\infty} (-1)^n q^{n(5n+3)/2} \quad \text{(B)}$$

$$(q^2;q^5)_\infty (q^3;q^5)_\infty (q^5;q^5)_\infty = \prod_{n\geq 0}\left\{(1+q^{5n+2})(1-q^{5n+3})(1-q^{5n+5})\right\} = \sum_{n=-\infty}^{\infty} (-1)^n q^{n(5n+1)/2} \quad \text{(C)}$$

In these cases, we have that $A \equiv A_0(\mathcal{C}, \delta) = 2\left(\frac{2}{a}\right)^{1/4}$ with $a = 3$ (A), 1, 1, 5 (B) and 5 (C), respectively. A comparison of these generating functions extending the visual projection of a correlation matrix onto the penguin image is provided below for reference with the OGFs labeled (A), (B) and (C) corresponding to the images in order from left to right.



Whenever (a, b) are integers such that $a \geq 1$ and $1 \leq b < a$, let

$$\mathcal{C}_{a,b}(q) := \sum_{n=-\infty}^{\infty} (-1)^n q^{\frac{n(an+b)}{2}}.$$

We compute the following special case values of $\mathrm{Corr}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$:

(a,b)	$(1,0)$	$(3,1)$	$(5,1)$	$(5,3)$	$(7,1)$	$(7,3)$	$(7,5)$
$\mathrm{Corr}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.76336	0.920081	0.0624965	0.672979	0.0108645	0.0374214	0.612813
(a,b)	$(11,1)$	$(11,3)$	$(11,5)$	$(11,7)$	$(11,9)$	$(13,1)$	$(13,3)$
$\mathrm{Corr}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.00158547	0.00264543	0.00470973	0.0202822	0.587536	0.000748935	0.00107876
(a,b)	$(13,5)$	$(13,7)$	$(13,9)$	$(13,11)$	$(17,15)$	$(23,21)$	$(29,27)$
$\mathrm{Corr}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.00182387	0.0046692	0.0188231	0.583074	0.58239	0.569502	0.56664

Lemma 6.11. *We have that the average order*

$$\mathbb{E}\left[\omega(\hat{p}(x)) + \omega(\hat{p}(-x))\right] \sim \left(\frac{a+4}{a^2}\right)(\log\log x + B),$$

where $B \approx 0.261497$ is the Mertens constant *(from Mertens' second theorem). That is, the average order of $\omega(n)$ over the two distinct integer-valued polynomials we get by expanding $p(\pm n)$ over all non-zero integers n is approximately a constant times $\mathbb{E}[\omega(x)]$ up to error terms that vanish as $x \to \infty$.*

The function $\omega(n)$ stays near its average order with a central normal tendency as [**?**, §1.7] [**?**, *cf.* §7.4]

$$\frac{1}{x} \times \#\left\{n \leq x : \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}} \leq z\right\} = \Phi(z) + o(1), z \in \mathbb{R}, \text{ as } x \to \infty.$$

We also have that uniformly for $0 < r \leq 1$

$$\#\{n \leq x : \omega(n) \leq r \log\log x\} \ll x(\log x)^{r-1-r\log r}, \text{ as } x \to \infty,$$

and that uniformly for $1 \leq r \leq R < 2$

$$\#\{n \leq x : \omega(n) \geq r \log\log x\} \ll_R x(\log x)^{r-1-r\log r}, \text{ as } x \to \infty.$$

Our intuition is hence to notice that $\omega(k)$ is so universally centered at its average order for almost every positive integer $k \geq 3$. Suppose that for $j, k \geq 16$ (so that the first n such that $\log\log(n) > 1$ is $n := \lceil e^e \rceil = 16$) we replace the terms in (48) involving $\omega(jk)$ with the corresponding average order formula from Lemma 6.11 evaluated at $x := jk$. We will denote this modified series by $\widehat{\mathrm{Corr}}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$. The following table (in comparison to the one given above) is suggestive of the regularity of these series and hence may be an approach towards obtaining tight bounds on the actual LGF correlation statistics:

(a,b)	$(1,0)$	$(3,1)$	$(5,1)$	$(5,3)$	$(7,1)$	$(7,3)$	$(7,5)$
$\widehat{\mathrm{Corr}}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.76336	0.920081	0.0624965	0.672979	0.0108645	0.0374214	0.612813
(a,b)	$(11,1)$	$(11,3)$	$(11,5)$	$(11,7)$	$(11,9)$	$(13,1)$	$(13,3)$
$\widehat{\mathrm{Corr}}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.00158547	0.00264543	0.00470973	0.0202822	0.587536	0.000748935	0.00107876
(a,b)	$(13,5)$	$(13,7)$	$(13,9)$	$(13,11)$	$(17,15)$	$(23,21)$	$(29,27)$
$\widehat{\mathrm{Corr}}_{\mathrm{LGF}}(\mathcal{C}_{a,b})$	0.00182387	0.0046692	0.0188231	0.583074	0.58239	0.569502	0.56664

Proof of Lemma 6.11. Suppose that we are evaluating the following sum:

$$E_{a,b}(x) = \frac{1}{\sqrt{x}} \times \sum_{\substack{n \geq 1 \\ \frac{n(an \pm b)}{2} \leq x}} \left[\omega\left(\frac{n(an+b)}{2}\right) + \omega\left(\frac{n(an-b)}{2}\right)\right].$$

We can perform a change of variable in the form of $v = \frac{n(an \pm b)}{2}$ so that $n = \frac{\sqrt{b^2+8av} \mp b}{2a}$ and $dn = \frac{2dv}{\sqrt{8av+b^2}}$. We know that the average order of the original function $\omega(n)$ is given by [8, §22.10]

$$\mathbb{E}[\omega(n)] = \frac{1}{x} \times \sum_{n \leq x} \omega(n) = \log\log n + B + o(1).$$

Then we have by the Abel summation formula, taking the summatory function $A(t) := t(\log\log t + B + o(1))$, that

$$\begin{aligned}
E_{a,b}(x) &= \frac{4}{\sqrt{x}} \times \sum_{v \leq x} \frac{\omega(v)}{\sqrt{8av+b^2}} \\
&\sim \frac{4}{\sqrt{x}} \left(\frac{(\log\log x + B)x}{\sqrt{8ax+b^2}} + \int_3^x \frac{4av(\log\log v + B)}{(8av+b^2)^{\frac{3}{2}}} dv \right) \\
&\sim \frac{1}{2}\sqrt{\frac{8}{a}}(\log\log x + B) + \frac{16a}{(8a)^{\frac{3}{2}}\sqrt{x}} \times \int_3^x \frac{\log\log v + B}{\sqrt{v}} dv \\
&= \frac{1}{2}\sqrt{\frac{8}{a}}(\log\log x + B) + \frac{32a}{(8a)^{\frac{3}{2}}\sqrt{x}} \left((\log\log x + B)\sqrt{x} - 2\,\mathrm{Ei}\left(\frac{\log x}{2}\right)\right) \\
&\sim \frac{1}{2}\sqrt{\frac{8}{a}}\left(1 + \frac{4}{a}\right)(\log\log x + B). \qquad (*)
\end{aligned}$$

In determining the main term in transition from the second to last equation above, we have used that

$$\log\log x - \log\log 2 + 1 - \frac{3\log x}{8} \leq \mathrm{Ei}\left(\frac{\log x}{2}\right) \leq \log\log x - \log\log 2 + 1 - \frac{3\log x}{8} + \frac{11(\log x)^2}{144}.$$

The main term for $\frac{2E_{a,b}(x)}{\sqrt{8a}}$ in $(*)$ above corresponds to the average order formula we seek to evaluate since $n \leq \frac{\sqrt{8ax+b^2} \mp b}{2a}$ so that the correct scalar multiple in front of the average sum is similar to $\frac{2}{\sqrt{8ax}}$. □

6.5 Conjectures on the formal algebraic structure to canonically best factorization theorems

6.5.1 An analysis of the algebraic properties in the motivating LGF coefficient cases

Given the way in which we have chosen to expand the factorizations of $L_f(q)$ as

$$(f * 1)(n) = [q^n]L_f(q) = \frac{1}{\mathcal{C}(q)} \times \sum_{n \geq 1} \sum_{k=1}^{n} s_{n,k}[\mathcal{C}] f(k) \cdot q^n,$$

the form of the invertible, lower triangular matrices with entries given by the $s_{n,k}[\mathcal{C}]$ are independent of any arithmetic f that defines these expansions. Moreover, these matrices are completely determined by the choice of the reciprocal generating function factor of $\mathcal{C}(q)$ subject only to the requirement that $\mathcal{C}(0) \neq 0$. It follows that taking an interpretation of the canonically "best possible" choice of this function being given by $\mathcal{C}(q) := (q;q)_\infty$, a criteria which we define qualitatively as inducing an unexpected, or particularly revealing substructure to the left-hand-side divisor sums, $f * 1$, is also independent of any fixed arithmetic f that defines $L_f(q)$.

Definition 6.12 (The Euler transform)**.** We observe a property called the *Euler transform* which nicely suggests motivation for why the partition function $p(n)$ arises so naturally in this class of LGF examples. Namely, we borrow the canonical integer sequence transformation identified by Bernstein and Sloane (circa 2002) called `EULER` [?]. It states that if two arithmetic functions a_n, b_n with a corresponding former OGF defined by $A(x) := \sum_{n \geq 1} a_n x^n$ are related by the identity

$$1 + \sum_{n \geq 1} b_n x^n = \prod_{i \geq 1} \frac{1}{(1-x^i)^{a_i}} \equiv \exp\left(\sum_{k \geq 1} \frac{A(x^k)}{k}\right),$$

then we can explicitly relate these sequence by introducing an intermediate divisor sum, denoted by c_n. In particular, if $c_n := \sum_{d|n} d \cdot a_d$, then we have that

$$a_n = \frac{1}{n} \times \sum_{d|n} c_d \mu\left(\frac{n}{d}\right), n \geq 1.$$

We know from elementary number theory due to Euler that the partition function $p(n)$ is related to the (ordinary) sum-of-divisors function, $\sigma(n) \equiv \sigma_1(n) := \sum_{d|n} d$, through the following recurrence relation:

$$np(n) = \sum_{0 \leq k < n} \sigma_1(n-k)p(k), n \geq 1.$$

Since $p(0) = 1$, the resulting ODE for the generating functions that relate these two sequences shows that

$$p(n) = [q^n] \exp\left(\sum_{k \geq 1} \frac{\sigma_1(k) q^k}{k}\right).$$

On the other hand, when we take the constant sequence $a_n \equiv 1, \forall n \geq 1$, the product expanded through the `EULER` transformation we defined above corresponds to the infinite q-Pochhammer function product, $(q;q)_\infty^{-1}$, which again generates $p(n)$ for all $n \geq 0$. Since $\sigma_{-\alpha}(n) = \sigma_\alpha(n) n^{-\alpha}$ for all $n \geq 1$ and any real parameter $\alpha \geq 0$, taking the exponential of the sum over the OGF $A(q)$ yields that

$$\prod_{i \geq 1} (1-q^i)^{-1} = \exp\left(\sum_{k \geq 1} \frac{q^k}{k(1-q^k)}\right) = \exp\left(\sum_{k \geq 1} \frac{\sigma_1(k) q^k}{k}\right).$$

Thus, we reason that the fundamental relation for $p(n)$ to the multiplicative divisor sums $\sigma_1(n)$ explains why the partition function arises here in the LGF case. We can look to the LGF special case for clues to see a good first order heuristic that we can use to measure how closely related the matrix and inverse matrix coefficients are for a fixed $\mathcal{D}$-convolution summation type. We clearly must define our metric to qunatify this heuristic so that is depends only on the kernel function $\mathcal{D}$, and the OGF $\mathcal{C}(q)$, and is always (of course) independent of f, g.

6.5.2 Partition theoretic conjectures based on explicit OGF series

Conjecture 6.13. *An optimal OGF $\mathcal{C}(q)$ that maximizes the correlation coefficients in (47), is given by*

$$\mathcal{C}(q) := \prod_{k \geq 1} \left(\sum_{n \geq 0} \mathcal{D}(n+k-1, k) q^n\right)^{-1}.$$

Conjecture 6.14. *The LGF OGF matchings we saw in Section 6.5.1 by applying the Euler transform of sequences suggests an optimal selection of $\mathcal{C}(q)$ satisfies the following expansions:*

$$\mathcal{C}(q) = \exp\left(-\sum_{n \geq 1} \sum_{k=1}^{n} k\mathcal{D}(n,k) \frac{q^n}{n}\right) = \prod_{n \geq 1} \left(1 + q \times \sum_{k=1}^{n} k\mathcal{D}(n,k)\right)^{-1}.$$

Conjecture 6.15 (Equivalence of problems)**.** *The cross-correlation statistic*

$$\mathrm{Corr}_1(\mathcal{C}, \mathcal{D}) := \sum_{n \geq 1} \frac{\frac{1}{n} \times \sum\limits_{1 \leq k \leq n} c_k(\mathcal{C}) \mathcal{D}^{-1}(n,k)}{\sqrt{\sum\limits_{1 \leq k \leq n} c_k(\mathcal{C})^2 \times \sum\limits_{1 \leq k \leq n} \mathcal{D}^{-1}(n,k)^2}},$$

is maximized (minimized) over all possible OGFs $\mathcal{C}(q)$ *if and only if*

$$\operatorname{Corr}_2(\mathcal{C},\mathcal{D}) := \sum_{n\geq 1} \frac{\frac{1}{n}\times \sum\limits_{1\leq k\leq n} p_k(\mathcal{C})\mathcal{D}(n,k)}{\sqrt{\sum\limits_{1\leq k\leq n} p_k(\mathcal{C})^2 \times \sum\limits_{1\leq k\leq n} \mathcal{D}(n,k)^2}},$$

is maximized (minimized) over all such OGFs.

6.5.3 Other conjectures

Remark 6.16. The values of certain signed sums are often modeled as a $\{\pm 1\}$-valued random walk on the integers whose height after the x^{th} step is taken to be approximately $M(x)$ where the probabilities of moving by ± 1 at any given step are randomized [**?**, **?**]. We know that the expectation of the absolute height at x of a prototypical random walk of this type is asymptotically $C\sqrt{x}$ for $C>0$ an absolute constant. Namely, suppose that $\{X_i\}_{i\geq 1}$ is a sequence of independent random variables defined such that $\mathbb{P}[X_i=-1]=\mathbb{P}[X_i=+1]=\frac{1}{2}$ for all $i\geq 1$. We can form the sums $Y_n := \sum_{i\leq n} X_i$ for $n\geq 1$. By computation we have that $\mathbb{E}[Y_n]=0$. At the same time, we can show that

$$\sigma_{Y_n} := \sqrt{\mathbb{E}[Y_n^2]-\mathbb{E}[Y_n]^2} = \sqrt{n}.$$

This follows since

$$Y_{n+1}^2 = (Y_n+X_n)^2 \implies \mathbb{E}[Y_{n+1}^2] = \mathbb{E}[Y_n^2]+1 \implies \mathbb{E}[Y_n^2]=n, \text{ for all } n\geq 1.$$

An interpretation of the combined first and second moment analysis is that we should expect the random walk modeled by Y_n to be approximately zero-valued most of the time but with an expected spread in actual values of as much as $\sqrt{n}$ [**?**]. The *law of the iterated logarithm* more precisely implies that $Y_n = O\left(\sqrt{n\log\log n}\right)$ for all sufficiently large n. We can also show that

$$\lim_{n\to\infty} \frac{\mathbb{E}|Y_n|}{\sqrt{n}} = \sqrt{\frac{2}{\pi}}.$$

Since the coefficients enumerated by a LGF for any function f are $f*1$, and we know by elementary number theory that $(\mu*1)(n)=(1*\mu)(n)=\delta_{n,1}$ (the multiplicative identity function with respect to Dirichlet convolution), we have that inversion of the convolution sums of this type is performed by Dirichlet convolution with $\mu(n)$ (*cf.* Möbius inversion). The summatory function, or partial sums of $\mu(n)$ are defined by $M(x) := \sum_{n\leq x}\mu(n)$ for any $x\geq 1$. The values of $M(x)$ are often modeled by a similar random walk whose values are $\{0,\pm 1\}$-valued according to the distribution of $\mu(n)\mapsto \pm 1$. This is often viewed as a case of the ± 1-valued random walk above with a different leading constant factor on the variance.

The Riemann Hypothesis (RH) is equivalent to proving that $M(x) = O\left(x^{\frac{1}{2}+\epsilon}\right)$ for all $0<\epsilon<\frac{1}{2}$. The dependence of the divisors $d|n$ over which we sum f to compute $(f*1)(n)$ at each $n\geq 1$ is deeply connected to the distribution of the primes. It stands to reason that the conventional interpretation of the primes as randomly determined in the sense of the ± 1-valued random walk model from above plays a pivotal role in the maximal correlation statistic that relates the kernel $\mathcal{D}^{-1}(n,k)=\mu\left(\frac{n}{k}\right)[k|n]_\delta$ to an optimal OGF, $\mathcal{C}(q)=(q;q)_\infty$ (as we have predicted it should be). The next conjecture cuts precisely to the crux of the matter with respect to why we seem to witness an optimal correlation statistic in the LGF case when $\mathcal{C}(q)$ satisfies $\rho_{\mathcal{C}}(n)^2 \asymp \sqrt{n}$.

Conjecture 6.17. *For a fixed lower triangular, invertible kernel function* $\mathcal{D}(n,k)$*, let*

$$M_{\mathcal{D}}(x) := \sum_{n\leq x} \mathcal{D}^{-1}(n,1), x\geq 1.$$

Suppose that

$$\delta = \left(\inf\left\{\rho>0 : M_{\mathcal{D}}(x) = O\left(x^{\rho+\varepsilon}\right), \forall \varepsilon>0\right\}\right)^2.$$

Then an optimal OGF $\mathcal{C}(q)$ *that witnesses the maximum possible value of* $\operatorname{Corr}(\mathcal{C},\mathcal{D})$ *satisfies* $\rho_{\mathcal{C}}(n)\asymp n^\delta$ *as* $n\to\infty$*. That is,* $\rho_{\mathcal{C}}(n)$ *is bounded above and below by absolute constant multiples of* n^δ *for all sufficiently large* n*.*