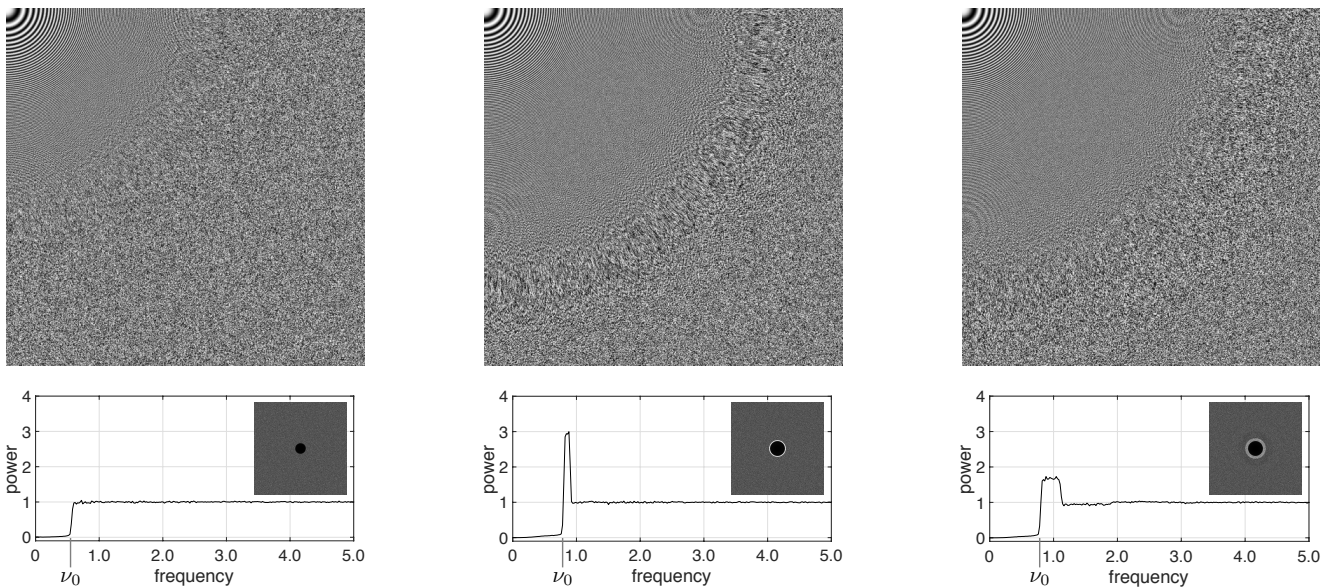


# A Comprehensive Theory and Variational Framework for Anti-aliasing Sampling Patterns

A. Cengiz Öztireli<sup>1,2</sup> 

<sup>1</sup>University of Cambridge

<sup>2</sup>Disney Research Studios



**Figure 1:** Anti-aliasing patterns, such as the step blue noise on the left, can generate images with clean low frequency content, and map higher frequencies to incoherent noise. The range of clean low frequencies (determined by  $\nu_0$  here) can be increased at the cost of introducing coherent colored noise for higher frequencies (middle, stair blue noise [KTBV16a]). Our approach can generate sampling patterns with any desired characteristics, e.g. (right) with minimal aliasing while keeping the same range of clean low frequencies (top: reconstructed zone plate images, bottom: 1D power spectra, insets: 2D power spectra).

## Abstract

In this paper, we provide a comprehensive theory of anti-aliasing sampling patterns that explains and revises known results, and introduce a variational optimization framework to generate point patterns with any desired power spectra and anti-aliasing properties. We start by deriving the exact spectral expression for expected error in reconstructing a function in terms of power spectra of sampling patterns, and analyzing how the shape of power spectra is related to anti-aliasing properties. Based on this analysis, we then formulate the problem of generating anti-aliasing sampling patterns as constrained variational optimization on power spectra. This allows us to not rely on any parametric form, and thus explore the whole space of realizable spectra. We show that the resulting optimized sampling patterns lead to reconstructions with less visible aliasing artifacts, while keeping low frequencies as clean as possible. Although we focus on image plane sampling, our theory and algorithms apply in any dimensions, and the variational optimization framework can be utilized in all problems where point pattern characteristics are given or optimized.

## CCS Concepts

• **Computing methodologies** → **Antialiasing**; **Rendering**; **Image processing**;

## 1. Introduction

Sampling patterns are fundamental for many applications in computer graphics such as imaging, rendering, geometry sampling, natural distribution modeling, among others. They are of particular importance for reconstructing functions from samples. Most real-world or synthesized functions are not band-limited, i.e. they contain frequencies higher than those that can be represented with a finite number of samples, inevitably leading to aliasing. The challenge is avoiding aliasing artifacts that show up as secondary structures that are not present in the original function, while having the lower frequency content cleanly reconstructed. An ideal anti-aliasing sampling pattern thus preserves lower frequencies by introducing minimal noise, and maps all higher frequencies that cannot be represented with the sample budget to incoherent noise, instead of visible artifacts [DW85, Coo86].

Many techniques have been proposed to generate sampling patterns for a larger range of cleanly represented frequencies, while avoiding aliasing artifacts as much as possible [Coo86, Uli88]. Recent techniques assume provided spectra or related statistics, and optimize locations of sampling points such that the resulting distributions have the given statistics [ZHWW12, OG12, HSD13, WPC\*14, AHD15, KTBV16a]. The challenge, however, is how to specify useful shapes for power spectra  $P(\nu)$  in the limited space of realizable spectra [UST06]. Recent works have focused on generating realizable spectra with certain properties beneficial for anti-aliasing [HSD13, KTBV16a] such as an as large as possible low-energy low-frequency region (given by  $\nu_0$ , two times the effective Nyquist frequency [HSD13], in Figure 1), while minimizing deviations from 1 for higher frequencies, with the intuition that the former will ensure clean low frequencies while the latter will lead to minimal aliasing.

These methods assume a parametric form for power spectra, and search in the parameter space to have the least energy in the low frequency region, and a flat high frequency region bounded from above. Such a sampling pattern generated by a state-of-the-art technique [KTBV16a] is shown in Figure 1, middle. The  $\nu_0$  is significantly larger than that of step blue noise (left), which manifests itself as a larger range of clean low frequencies in the reconstructed zone plate image. However, this comes at the cost of a flat peak in the power spectrum, introducing artifacts in the zone plate image for middle frequencies. In general, assuming a given parametric form limits power spectra, leading to suboptimal anti-aliasing properties.

In this paper, we introduce a non-parametric power spectra optimization framework and show how it can be used to generate point patterns with anti-aliasing properties based on a comprehensive theoretical analysis. We can then generate point patterns with any desired properties such as the ones in Figure 1, spanning the whole space of possible point patterns. In order to formulate the corresponding optimization problem, we first prove an analytic form for the spectrum of expected error introduced by sampling. Based on this formula for error, we show how existing patterns improve anti-aliasing, and provide new theoretical results and insights. These are then translated into constraints and energies for formulating a constrained variational optimization on power spectra of point patterns. We show that careful selection of constraints and energies

to minimize lead to sampling patterns with improved anti-aliasing properties. A sampling pattern generated by the proposed technique is shown in Figure 1, right. We get the same  $\nu_0$  and thus range of noise-free lower frequencies as for the result of Kailkhura et al. [KTBV16a], while still mapping all higher frequencies to almost white noise, as can be seen in the zone plate test image. The resulting power spectrum arises from our formulation of the optimization, without explicitly specifying its form.

In summary, we have the following main contributions:

- A theory of anti-aliasing with exact expressions for expected error spectrum in any dimension. This allows us to analyze desirable properties for power spectra of point patterns for anti-aliasing.
- A new formulation of the problem of generating realizable spectral or spatial characteristics of point patterns based on variational optimization.
- Sampling patterns optimized for anti-aliasing with practical improvements over state-of-the-art patterns.

## 2. Related Work

Aliasing is a fundamental problem when reconstructing or synthesizing functions, e.g. images, with samples, as they are typically not band-limited and we always have a finite budget of samples. It is well-known that regular sampling leads to structured aliasing, which introduces visually distracting extra structures. A main observation is that by injecting randomness into point distributions while satisfying certain properties, structured artifacts can be replaced with noise that is potentially visually less distracting [DW85, Coo86]. With such random distributions, it is important that high frequencies that cannot be represented with the sample budget are mapped to as incoherent as possible noise, ideally white noise to avoid any extra patterns in the reconstructed image, while keeping the important low frequency content clean.

Such sampling patterns are typically called *blue noise* in computer graphics. Blue noise patterns are characterized by a low energy power spectrum  $P(\nu)$  for  $\nu < \nu_0$ , and a flat spectrum with  $P(\nu) \approx 1$  for  $\nu > \nu_0$  [Yel83, Mit91]. Many methods have been proposed to generate point patterns with power spectra that exhibit variations of such properties. Earlier methods propose algorithms that impose certain constraints on the generated random point distributions. Dart throwing [Coo86] (also known as simple sequential inhibition and random sequential adsorption [IPSS08]) generates distributions where points are randomly placed in space with the constraint that no two points are closer to each other than a certain distance. This algorithm and the resulting distributions have been widely used and extended in many ways in the last decades (e.g. [DH06, Bri07, Wei08, Wei10, EMP\*12, EPM\*14, Yuk15, KTBV16b]). Other works have investigated utilizing alternative algorithms for improved characteristics for certain applications [KCODL06, Ost07, IPSS08, BSD09, SGBW10, Fat11, SHD11, XLGG11, CYC\*12, dGBOD12, JZW\*15, MEA\*18]. The resulting point distributions are then analyzed by computing characteristics such as power spectrum, or statistics from stochastic point processes [Mit87, LD08, WW11, OG12, HSD13], to understand their utility in practice.

A main limitation of the mentioned works for point pattern generation, however, is that the algorithm dictates the characteristics of the generated point patterns. Instead, a recent body of works propose to generate point distributions with statistics matching given ones [ZHWW12, OG12, HSD13, WPC\*14, AHD15, KTBV16a]. Once a statistic, such as the power spectrum, is defined, these methods run a routine to place sampling points such that the final configuration leads to the desired form for the statistic. With this approach, Heck et al. [HSD13] could generate point distributions with the step blue noise spectrum for the first time (Figure 1, left). However, they have also observed that such a form for the spectrum is only possible for quite low values of  $\nu_0$ , leading to noisy lower frequencies for sampled images. In general, the subspace of realizable power spectra is restricted, with the necessary conditions that both power spectrum and pair correlation function, which is related to power spectrum with a spectral transform, should be non-negative [UST06]. Hence, a fundamental challenge is defining realizable forms for power spectra with desirable properties.

This challenge has been addressed by defining parametric forms for the power spectrum in recent works [HSD13, KTBV16a]. The idea is then to search over the free parameters to get realizable power spectra with anti-aliasing properties. Heck et al. [HSD13] define the *single-peak blue noise*, where a Gaussian is placed at around  $\nu_0$  to trade off energy for  $\nu < \nu_0$  against the maximum value  $m$  of the power spectrum. The standard deviation, and magnitude of the Gaussian can then be altered to get realizable power spectra. Kailkhura et al. [KTBV16a] have recently proposed a new parametrized family, *stair blue noise*, where the peak is replaced with a raised flat region of a certain width starting at  $\nu_0$ , as in Figure 1, middle. The free parameters in this case are  $\nu_0$ , and the width and height of the raised flat region. By a guided search over these parameters, spectra with a lower  $m$  than single-peak blue noise can be obtained. However, for both methods, due to the assumed parametric forms, the families of power spectra considered are rather limited, and the exact effect of the parameters on aliasing is not clear. In contrast, we do not assume any particular form for power spectra and instead formulate the problem of generating desirable and realizable spectra as constraint optimization with a variational formulation.

### 3. Background and Preliminaries

We utilize the theory of stochastic point processes [MW04, IPSS08] to understand anti-aliasing properties of point patterns. Stochastic point processes provide a principled approach for analyzing point patterns [MW04, IPSS08]. A point process is defined as the generating process for multiple point distributions sharing certain characteristics. Hence, each distribution can be considered as a realization of an underlying point process (we use the term *point pattern* for families of point distributions sharing characteristics).

We can explain a point process with joint probabilities of having points at certain locations in space. Such probabilities are expressed in terms of *product densities*. For our application of point patterns with optimal power spectra for anti-aliasing, it is sufficient to consider first and second order product densities, as they uniquely determine the power spectrum of a point process. First order product density is given by  $\rho^{(1)}(\mathbf{x})d\mathbf{x} = p(\mathbf{x})$ , where  $p(\mathbf{x})$  is the probability

of having a point generated by the point process  $\mathcal{P}$  in the set  $d\mathbf{x}$  of infinitesimal volume, and intuitively measures expected number of points around  $\mathbf{x}$ , i.e. local density. Similarly, second order product density  $\rho^{(2)}$  is defined in terms of the joint probability  $p(\mathbf{x}, \mathbf{y})$  of having points  $\mathbf{x}$  and  $\mathbf{y}$  in the sets  $d\mathbf{x}$  and  $d\mathbf{y}$  simultaneously,  $\rho^{(2)}(\mathbf{x}, \mathbf{y})d\mathbf{x}d\mathbf{y} = p(\mathbf{x}, \mathbf{y})$ . It describes how points are arranged in space, and is fundamentally related to the power spectrum of  $\mathcal{P}$ .

As in previous works [DW85, HSD13, KTBV16a] on anti-aliasing, we will assume that no information is given on the function to be represented, and hence consider unadaptive point patterns. These patterns are generated by stationary and isotropic point processes, where the characteristics of the generated point distributions are translation invariant, or translation and rotation invariant, respectively [OG12]. For both cases,  $\rho^{(1)}$  reduces to a constant number,  $\lambda$ , which measures the expected number of points in any given volume,  $\lambda = \frac{\mathbb{E}_{\mathcal{P}}[n(\mathcal{V})]}{|\mathcal{V}|}$ , where  $\mathbb{E}_{\mathcal{P}}$  denotes expectation over different distributions generated by the point process  $\mathcal{P}$ ,  $n(\mathcal{V})$  is the random number of points that fall into the set  $\mathcal{V}$ , and  $|\mathcal{V}|$  is its volume. For stationary point processes, second order product density becomes a function of the difference vector between point locations  $\rho^{(2)}(\mathbf{x}, \mathbf{y}) = \rho^{(2)}(\mathbf{x} - \mathbf{y})$ , which can be expressed in terms of the normalized *pair correlation function (PCF)*  $g$  as  $\rho^{(2)}(\mathbf{x} - \mathbf{y}) = \lambda^2 g(\mathbf{x} - \mathbf{y})$ . For isotropic point processes, PCF further simplifies and becomes a function of the distance between point locations  $g(\mathbf{x} - \mathbf{y}) = g(\|\mathbf{x} - \mathbf{y}\|)$ . Below we will first consider stationary point processes and the associated derivations, which we will specialize to isotropic processes in the next sections.

#### PCF as a Distribution

The intuition behind PCF is that it can be estimated as a probability distribution of difference vectors (for stationary point processes), or distances (for isotropic point processes) between points. This is possible due to the fundamental Campbell's theorem [IPSS08] that relates sums of functions at sample points to integrals of those functions. For simplicity of the expressions, we assume a toroidal domain  $\mathcal{V}$  with unit volume as the sampled domain (e.g. the image plane). Utilizing Campbell's theorem, it is possible to derive the following expression for PCF of stationary processes (Appendix A of the supplementary material)

$$g(\mathbf{r}) = \frac{1}{\lambda^2} \mathbb{E}_{\mathcal{P}} \left[ \sum_{j \neq k} \delta(\mathbf{r} - \mathbf{r}_{jk}) \right], \quad (1)$$

where  $\delta$  is the Dirac delta, and we defined  $\mathbf{r} = \mathbf{x} - \mathbf{y}$ ,  $\mathbf{r}_{jk} = \mathbf{x}_j - \mathbf{x}_k$ . Note that the  $\mathbf{x}_k$ 's are from a particular distribution generated by the point process  $\mathcal{P}$ , and the expectation is over all such distributions. This expression clearly shows that PCF is simply a normalized distribution of difference vectors  $\mathbf{r}_{jk}$ .

#### Power Spectrum and PCF

Power spectrum of a point process is defined in terms of the Fourier transform  $S(\mathbf{v}) = \mathcal{F}[s(\mathbf{x})](\mathbf{v})$  of the function  $s(\mathbf{x}) = \sum_j \delta(\mathbf{x} - \mathbf{x}_j)$  as follows [HSD13]

$$P(\mathbf{v}) = \frac{1}{\lambda} \mathbb{E}_{\mathcal{P}} \left[ \overline{S(\mathbf{v})} S(\mathbf{v}) \right] = \frac{1}{\lambda} \mathbb{E}_{\mathcal{P}} \left[ \sum_{jk} e^{-2\pi i \mathbf{v}^T \mathbf{r}_{jk}} \right], \quad (2)$$



with  $\overline{(\cdot)}$  denoting complex conjugate. Power spectrum  $P$  thus lacks the phase of the Fourier transform and hence is translation invariant, only depending on the difference vectors  $\mathbf{r}_{jk}$ . Equations 1 and 2 suggest that  $P$  and  $g$  are related by a Fourier transform. Indeed, denoting the Fourier transform of  $g$  with  $G$ , it is possible to derive the following relation between them (Appendix B of the supplementary material)

$$P(\mathbf{v}) = \lambda G(\mathbf{v}) + 1. \quad (3)$$

In order to state properties of power spectra for anti-aliasing, we will work with a slightly modified form of Equation 3, where we rewrite the relation between  $P$  and  $g$  in terms a function  $u$  we define, and its Fourier transform  $U$ , as follows

$$g(\mathbf{r}) = u(\mathbf{r})/\lambda + 1, \quad P(\mathbf{v}) = U(\mathbf{v}) + 1 + \lambda\delta(\mathbf{v}). \quad (4)$$

### Conditions for Realizable Power Spectra

Power spectrum is non-negative by definition (Equation 2), and this is also true for PCF as a distribution of difference vectors (Equation 1). Hence, two necessary conditions for a valid power spectrum of a point process are

$$g(\mathbf{r}) \geq 0, \quad P(\mathbf{v}) \geq 0. \quad (5)$$

It is still an open question whether these are also sufficient conditions, but no counterexamples have been shown in statistics and physics (e.g. [TS02]), and these conditions have been successfully used to generate realizable power spectra in previous works [HSD13, KTBV16a] and also in this paper. We thus assume that these are sufficient conditions for realizable power spectra.

### Error in Sampling a Function

Sampling a function  $t$  with a point distribution generated by a point process  $\mathcal{P}$  can be written as  $s(\mathbf{x})t(\mathbf{x})$  in the spatial domain, or as  $[S * T](\mathbf{v})$  in the frequency domain, where  $T$  is the Fourier transform of  $t$ , and  $*$  denotes convolution. This sampled representation introduces an error. In order to analyze magnitude and distribution of error, the expected power spectrum of error needs to be computed [DW85, HSD13]

$$E(\mathbf{v}) = \mathbb{E}_{\mathcal{P}} \left[ \left| [S * T](\mathbf{v})/\lambda - T(\mathbf{v}) \right|^2 \right], \quad (6)$$

where  $|\cdot|$  denotes magnitude of a complex number, and the sampled representation is divided by  $\lambda$  to normalize the energy of the sampled function [HSD13], or equivalently to have an unbiased estimator since  $\mathbb{E}_{\mathcal{P}}[s(\mathbf{x})t(\mathbf{x})/\lambda] = t(\mathbf{x})$  (by applying Equation 3 in Appendix A of the supplementary material). We need to relate the error  $E(\mathbf{v})$  to  $P(\mathbf{v})$  in order to derive desired properties for this statistic  $P(\mathbf{v})$ , which we elaborate on in the next section.

### Relation to integration

In this work, we are interested in error when representing a function with samples. This is fundamentally different from the error introduced by numerically integrating a function by summing the sample values [Dur11, SK13, PSC\*15, Ö16]. For the case of images, we can think of the sampling, filtering, and resampling of an image as performing *local* integration around each pixel center. To understand aliasing, we need to analyze the *distribution* of these

errors of integral estimates at all pixels. Indeed, we are interested in the spectrum of error that encodes this distribution. This is in contrast with analyzing error in a single integral estimate. However, some insights from integral estimation carry over to the anti-aliasing problem, as we will see in the next section.

## 4. Theoretical Analysis of Sampling Error

The error spectrum  $E(\mathbf{v})$  provides how much error we get at each frequency. In general, we need to have  $E(\mathbf{v})$  as low as possible at each  $\mathbf{v}$ , and especially for low  $\mathbf{v}$ . For anti-aliasing, we need to additionally have an as uniform as possible  $E(\mathbf{v})$  to get incoherent noise instead of colored noise [DW85, HSD13]. It is, however, not clear how these are exactly related to the shape of the power spectrum  $P(\mathbf{v})$  of a sampling pattern. We need this relation to be able to formulate the constraints and energies for our variational formulation of optimized sampling patterns for anti-aliasing.

### 4.1. Spectra of Error and Sampling Patterns

So far, relating  $E(\mathbf{v})$  to  $P(\mathbf{v})$  has only been possible for a constant function  $t(\mathbf{x}) = c$  [DW85], or upper bounds could be derived for a sinusoidal wave [HSD13]. In this section, we show that it is possible to derive an exact relation between  $E(\mathbf{v})$  and  $P(\mathbf{v})$  for an arbitrary  $t(\mathbf{x})$ , by utilizing the theory of point processes. This leads to theoretical justifications of criteria used for  $P(\mathbf{v})$  in the literature, and to novel theoretical results and insights.

We start by expanding the expression for the error spectrum in Equation 6 (we drop  $\mathbf{v}$  for brevity)

$$\begin{aligned} E &= \frac{1}{\lambda^2} \mathbb{E}_{\mathcal{P}} \left[ |S * T|^2 \right] + \mathbb{E}_{\mathcal{P}} \left[ |T|^2 \right] \\ &\quad - \frac{1}{\lambda} \mathbb{E}_{\mathcal{P}} \left[ \overline{(S * T)} T \right] - \frac{1}{\lambda} \mathbb{E}_{\mathcal{P}} \left[ (S * T) \overline{T} \right] \\ &= \frac{1}{\lambda^2} \mathbb{E}_{\mathcal{P}} \left[ |S * T|^2 \right] + |T|^2 - \frac{2}{\lambda} \Re \left\{ \mathbb{E}_{\mathcal{P}} [S * T] \overline{T} \right\}, \end{aligned} \quad (7)$$

where  $\Re(\cdot)$  gives the real part of a complex number. The critical part of the proof is deriving the forms of these expected values. We show in Appendix A that this can be achieved by starting from Campbell's theorem (as defined in Appendix A of the supplementary material). The final form of the power spectrum of error is then

$$E(\mathbf{v}) = \frac{1}{\lambda} [P_t * (U + 1)](\mathbf{v}). \quad (8)$$

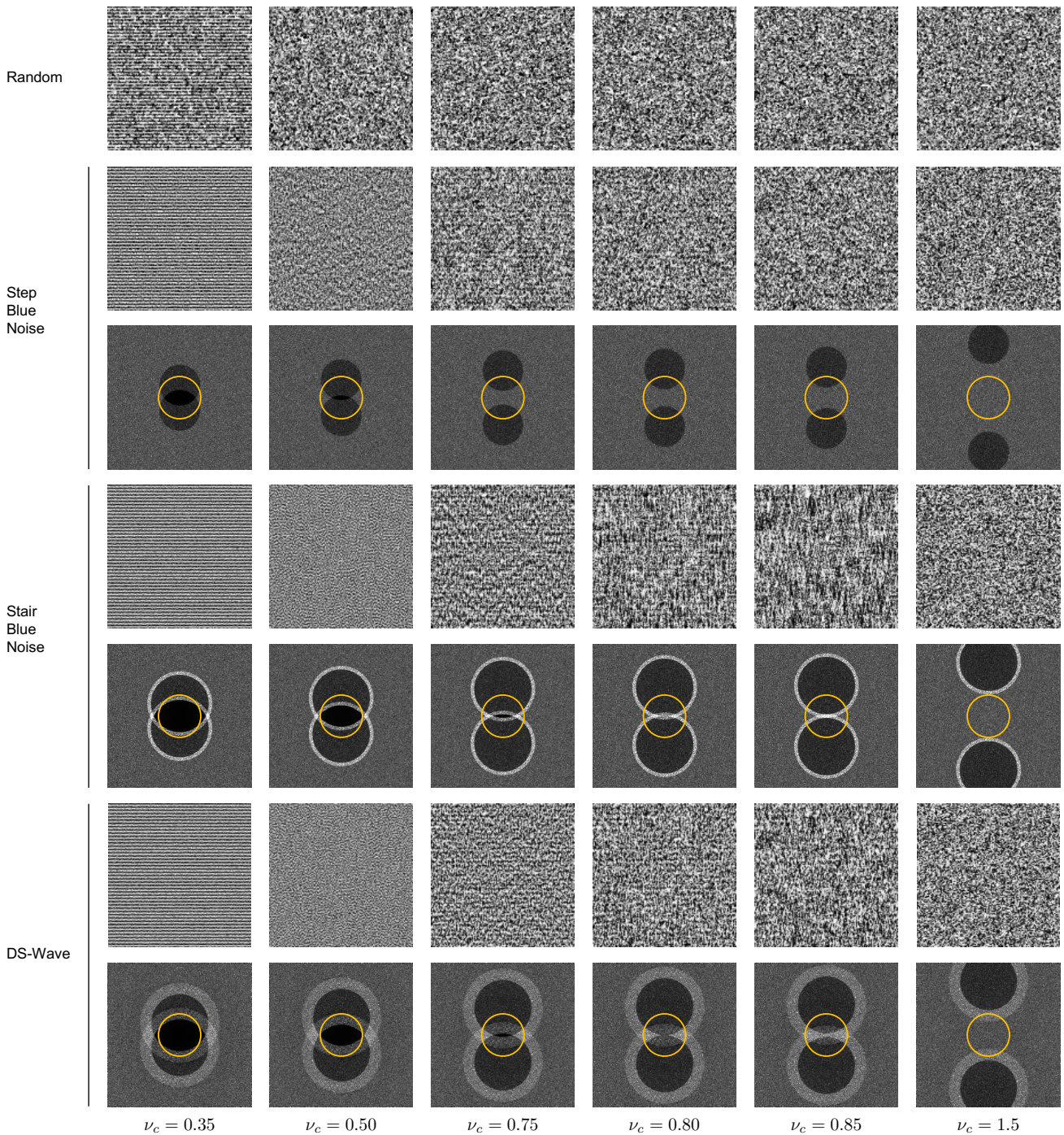
Here,  $P_t = |T|^2$  is the power spectrum of the function  $t$ .

### Remarks

This expression immediately reveals several interesting properties of error when sampling a function.

- The error is independent of the phase of  $T(\mathbf{v})$ . This is expected as the sampling patterns considered are translation invariant.
- It implies that error *can* decrease as  $O(\lambda^{-1})$  for any function  $t$ , as observed for a sinusoidal wave previously [HSD13] (recall that  $\lambda$  is proportional to the number of points). However, at the same time, the difference vectors  $\mathbf{r}_{jk}$  become smaller for higher number of points, leading to a compression of the domain of  $g(\mathbf{r})$





**Figure 2:** Colored noise leads to visible secondary structures that are not present in original images. The function  $\cos(2\pi\sqrt{\lambda}\nu_c y)$  is sampled with different sampling patterns (with  $\lambda = 128^2$ , and one sample per pixel). We show the resulting images as well as the error  $E(\mathbf{v})$ , with the orange circle marking the region of representable frequencies after reconstruction and resampling to the pixel grid. Stair blue noise reduces noise levels as compared to step blue noise for lower  $\nu_c = 0.35, 0.50$ , but leads to aliasing artifacts for higher  $\nu_c = 0.75, 0.80, 0.85$  due to the fluctuations it introduces to  $E(\mathbf{v})$  for the representable low frequencies within the orange circles. The proposed ds-wave sampling results in less or equivalent noise levels for all  $\nu_c$  as compared to other patterns, and less aliasing for higher  $\nu_c$  as compared to the state-of-the-art stair blue noise [KTBV16a] (both patterns are with  $\nu_0 = 0.8$ ) due to lower peaks in  $E(\mathbf{v})$ . For very high  $\nu_c = 1.5$ , all patterns have noise levels similar to random sampling.

(Equation 1), and hence an expansion of that of  $U(\mathbf{v})$ , as they are related via a Fourier transform. Thus, the final convergence rate depends on  $P_t(\mathbf{v})$  and  $U(\mathbf{v})$ .

- The only pattern that gives a constant spectrum is random sampling (Poisson point process) with  $U(\mathbf{v}) = 0$ . In this case, we get  $E(\mathbf{v}) = \frac{1}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v}) d\mathbf{v}$ , which leads to equally noisy frequencies and hence perfectly incoherent white noise.
- The DC component of sampling error given by  $E(\mathbf{0}) = \frac{1}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v})(U(\mathbf{v}) + 1) d\mathbf{v}$  is exactly the variance of the numerical estimator  $\frac{1}{\lambda} \sum_j t(\mathbf{x}_j)$  for the integral  $\frac{1}{|\mathcal{V}|} \int_{\mathcal{V}} t(\mathbf{x}) d\mathbf{x}$  [PSC\*15, Ö16]. For the stationary point processes we consider, bias vanishes and hence this variance is equal to the expected error of the numerical integral estimator [Ö16].

In practice, a function (e.g. image) sampled with an anti-aliasing point pattern is then resampled to a regular grid after low-pass filtering. This can be written as  $s_{REG}(k * (st))$  (dropping  $\mathbf{x}$  for brevity), where  $k$  is a low-pass filter such as Gaussian, and the points in  $s_{REG}$  are regularly distributed on a grid of e.g. pixel centers. This resampled function has the Fourier transform  $S_{REG} * (K(S * T))$  with  $S_{REG}$  and  $K$  the Fourier transforms of  $s_{REG}$  and  $k$ , respectively. As  $S_{REG}$  is an impulse train, the result of this convolution is repeating the same function, assuming  $K$  avoids any overlap between aliases. Thus, only the central part around zero frequency, cut out by the filter  $K$ , is relevant. The expected error (Equation 6) then becomes

$$\mathbb{E}_{\mathcal{P}} \left[ |K[S * T]/\lambda - KT|^2 \right] = |K|^2 E. \quad (9)$$

Hence, we can consider the low frequency region of  $E$  implied by  $K$  for most practical applications.

## 4.2. Analysis of Anti-aliasing Properties

The derived relation between  $E(\mathbf{v})$  and  $P(\mathbf{v})$  allows us to perform a theoretical analysis of error in terms of the characteristics of the power spectrum. There are established characteristics for the power spectra  $P(\mathbf{v})$  of anti-aliasing point patterns in the literature. These follow certain intuitions and have indeed been effective in practice. However, how such characteristics exactly affect aliasing, and how they can be improved, could not be analyzed since the relation between  $P(\mathbf{v})$  and  $E(\mathbf{v})$  was not known [HSD13].

There are two considerations for the error: 1) it should be low, 2) it should be as constant as possible, leading to white noise. The latter ensures that additional visual structures will not appear due to colored noise. We want to analyze how  $U(\mathbf{v})$  and thus  $P(\mathbf{v})$  should be shaped to achieve such a spectral profile for noise. For brevity, in the rest of the paper, we set  $P(\mathbf{v}) = U(\mathbf{v}) + 1$ , ignoring the Dirac delta at zero that does not contribute to  $E(\mathbf{v})$ .

### Low energy for low frequencies

A fundamental property of anti-aliasing patterns such as blue noise patterns is that there should be a low energy low frequency region [Mit91, HSD13, KTBV16a], i.e.  $P(\mathbf{v})$  should be low and ideally zero for  $\|\mathbf{v}\| < v_0$ . This property is meant to limit the amount of noise  $E(\mathbf{v})$  at lower frequencies. By expanding the convolution in Equation 8, it can be easily shown for step blue-noise (Figure 1,

left) where  $P(\mathbf{v}) = U(\mathbf{v}) + 1 = 0$  for  $\|\mathbf{v}\| < v_0$ , and 1 otherwise, that

$$\begin{aligned} E_{STEP}(\mathbf{v}) &= \frac{1}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v} - \mathbf{v}') P_{STEP}(\mathbf{v}') d\mathbf{v}' \\ &= \frac{1}{\lambda} \int_{\mathcal{D}_{v_0}^c} P_t(\mathbf{v} - \mathbf{v}') d\mathbf{v}' \\ &\leq \frac{1}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v} - \mathbf{v}') d\mathbf{v}' = E_{RND}(\mathbf{v}), \end{aligned} \quad (10)$$

where  $\mathcal{D}_{v_0}$  is the  $d$ -dimensional disk of radius  $v_0$ ,  $(\cdot)^c$  denotes the complement of a set, and  $E_{STEP}$  and  $E_{RND}$  are the errors when using step blue noise, and random sampling, respectively. In particular, for a band-limited function with  $P_t(\mathbf{v}) = 0$  for  $\|\mathbf{v}\| > v_0$ ,  $E_{STEP} = 0$ . However, in general, the difference between  $E_{STEP}$  and  $E_{RND}$  may not be very large especially when the function  $P_t(\mathbf{v})$  has significant energy at higher frequencies. This can be seen in Figure 2, where we show examples of sampled images of a cosine wave  $\cos(2\pi\sqrt{\lambda}v_c y)$  of different frequencies  $v_c$ , and the corresponding  $E(\mathbf{v})$ , for different sampling patterns. For high frequencies such as  $v_c = 1.5$ , the low frequency region (implied by the reconstruction kernel  $K$ , marked with orange circles in the figure) of  $E(\mathbf{v})$  for step blue noise contains as much energy as for random sampling, leading to similar levels of error in the sampled images.

In practice, having a large  $v_0$  is still very important even when sampling non-band-limited functions, due to the stationarity of the point patterns considered (Section 3). As the patterns are translation invariant, each local patch of the function  $t$  is sampled with a point distribution of the same characteristics. Hence, the same analysis can be carried out for each patch. The visual quality especially for smoother patches, where noise is visually very distracting, will thus be improved significantly by using a step-like profile. This is illustrated in Figure 2 for  $v_c = 0.35$ , where stair blue noise and ds-wave sampling (a variation of our sampling patterns as we will discuss in Section 5) with a higher  $v_0$  than step blue noise and random sampling, result in much cleaner image content.

### Limiting the maximum of $P$

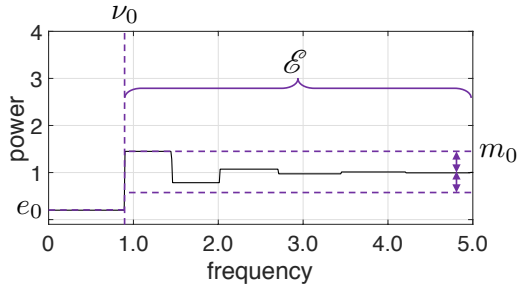
Another fundamental property utilized in the literature is that the maximum  $m$  of  $P$  should be limited [HSD13, KTBV16a]. The intuition is that this will also limit the magnitude of and fluctuations in error. Indeed, we can easily show that

$$E(\mathbf{v}) \leq \frac{m}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v} - \mathbf{v}') d\mathbf{v}' = \frac{m}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v}) d\mathbf{v}. \quad (11)$$

Hence, normalized by the total energy of  $T(\mathbf{v})$ , the error in this case is bounded by  $m/\lambda$  at every frequency  $\mathbf{v}$ .

This global maximum is also very important for limiting fluctuations in  $E(\mathbf{v})$ , i.e. avoiding colored noise. Example  $E(\mathbf{v})$ 's where such maxima add up to generate significant fluctuations in error for low frequencies are shown in Figure 2, stair blue noise sampling with  $v_c = 0.75 - 0.85$ . In this case,  $P_t(v_1, v_2) = \alpha[\delta(v_2 - \sqrt{\lambda}v_c) + \delta(v_2 + \sqrt{\lambda}v_c)]\delta(v_1)$  for a constant  $\alpha$ , and hence the ratio of error to the total energy of the function  $t$  is  $E(\mathbf{v}) / \int_{-\infty}^{\infty} P_t(\mathbf{v}) d\mathbf{v} = \frac{1}{2\lambda} [P(v_1 - \sqrt{\lambda}v_c) + P(v_1 + \sqrt{\lambda}v_c)]$ . In general for any function  $t$ , this ratio can fluctuate between 0 and  $m/\lambda$  at different frequencies, significantly disturbing the noise profile if the maximum  $m$  is high.





**Figure 3:** Realizable power spectra with anti-aliasing properties can be obtained by constrained variational optimization with constraints on the maximum value  $e_0$  of the low frequency region  $\mathbf{v} < \mathbf{v}_0$ , deviation from 1 for the high frequency region  $\mathbf{v} > \mathbf{v}_0$  that implies the constraint  $m_0$  on the maximum value of the spectrum, and an energy functional  $\mathcal{E}$  that controls the shape of the high frequency region.

Such colored noise manifests itself as visually distinguishable secondary patterns in sampled images, as can be seen in the image reconstructions for stair blue noise with  $\mathbf{v}_c = 0.75 - 0.85$  in the figure, instead of white noise without a clear structure.

#### Minimizing local maxima of $P$

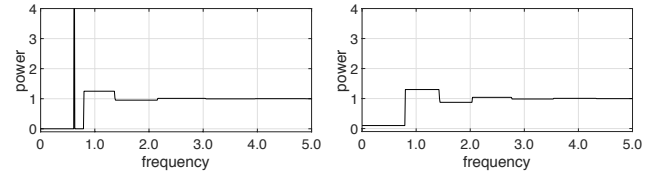
Due to the constraints on  $P(\mathbf{v})$  (Section 3), for point patterns with a larger  $\mathbf{v}_0$ ,  $P(\mathbf{v})$  inevitably contains local maxima of decaying magnitude (as we will illustrate in Section 5). Apart from limiting  $m$ , which determines the first maximum in  $P(\mathbf{v})$ , avoiding further local maxima is also beneficial, as these peaks can similarly sum up to cause further fluctuations in  $E(\mathbf{v})$  due to the convolution in Equation 8, albeit all smaller than  $m$  as we illustrate in Figure 2. We will explore how we can shape the peaks such that we get an as small as possible global maximum  $m$  and local maxima, while ensuring a certain  $\mathbf{v}_0$ , by translating these into energies and constraints in a variational optimization based formulation for  $P(\mathbf{v})$  in the next section.

### 5. Optimized Anti-aliasing Patterns

The characteristics for  $P$  as elaborated on in the last section can be imposed in addition to the realizability conditions (Equation 5), to obtain optimal sampling patterns with respect to these criteria. In this section, we formulate the associated variational optimization problem. This will allow us to synthesize optimal distributions with respect to the considered characteristics with numerical solution methods.

#### 5.1. Sampling as Constrained Variational Optimization

We start by making the effect of density  $\lambda$  on the problem explicit, and factor it out from the optimization. This can be achieved by working with normalized spatial and frequency coordinates. We start by defining  $f(\mathbf{r}) = u(\mathbf{r}/\lambda^{1/d})/\lambda$ . By the scaling property of Fourier transform, we can write  $F(\mathbf{v}) = U(\mathbf{v}\lambda^{1/d})$ . Substituting these into the expressions for  $g$  and  $P$  (Equation 4), we get



**Figure 4:** Power spectra generated with an integral-based low frequency constraint can lead to spikes for low frequencies (left). Directly limiting power for low frequencies ensures low aliasing (right).

$g(\mathbf{r}/\lambda^{1/d}) = f(\mathbf{r}) + 1$ , and  $P(\mathbf{v}\lambda^{1/d}) = F(\mathbf{v}) + 1$  (ignoring  $\delta(\mathbf{v})$  as before, as it does not contribute to  $E(\mathbf{v})$ ). Then, in normalized coordinates, we can write the conditions  $g(\mathbf{r}/\lambda^{1/d}) \geq 0$  and  $P(\mathbf{v}\lambda^{1/d}) \geq 0$  as

$$f(\mathbf{r}) + 1 \geq 0, \quad F(\mathbf{v}) + 1 \geq 0. \quad (12)$$

Thus, the constraints become independent of the intensity  $\lambda$  of the point process. We will work with  $g$  and  $P$  in normalized coordinates unless stated otherwise, and set  $g(\mathbf{r}) = f(\mathbf{r}) + 1$ , and  $P(\mathbf{v}) = F(\mathbf{v}) + 1$ . Absolute spatial coordinates are thus given by multiplying the reported  $\mathbf{r}$  with  $1/\lambda^{1/d}$ , and absolute frequencies by multiplying the reported  $\mathbf{v}$  with  $\lambda^{1/d}$ .

Although the complete analysis in the rest of the paper can be carried out for stationary point processes in  $\mathbf{R}^d$ , we will consider the important case of image sampling with non-adaptive anti-aliasing distributions as in previous works [DW85, HSD13, KTBV16a]. This implies that the point processes considered are isotropic, generating rotation and translation invariant distributions. In this case,  $g$  and thus  $P$  is radially symmetric such that  $g(\mathbf{r}) = g(\|\mathbf{r}\|)$ ,  $P(\mathbf{v}) = P(\|\mathbf{v}\|) = P(v)$ , and all Fourier transforms in the definitions above turn into Hankel transforms  $\mathcal{H}$  (please see Appendix A of Heck et al. [HSD13] for the definition and properties of Hankel transforms). In particular, we have  $F = \mathcal{H}[f]$ , or equivalently  $f = \mathcal{H}[F]$ . The Hankel transform is defined for any dimension. For our case of sampling the image plane, we use the Hankel transform for  $d = 2$  dimensions.

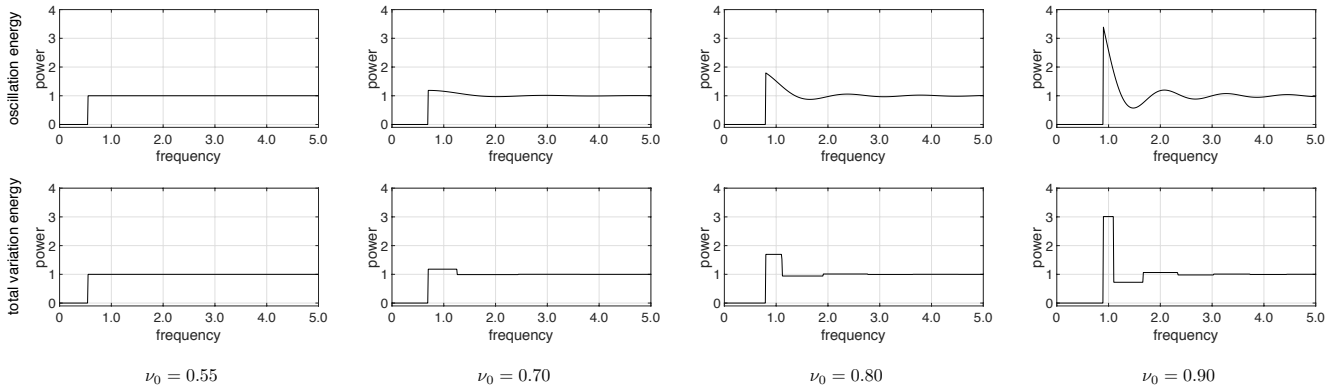
Hence, the problem becomes finding a 1D function  $F$  with the above non-negativity constraints in Equation 12, and additional properties we impose. These properties can either be set as hard inequality constraints  $\mathcal{C}[F(\mathbf{v})] \geq 0$ , or energies that we minimize for. We can thus formulate the following constrained variational minimization problem on  $F$ , to find a realizable and desirable  $P$

$$\min \mathcal{E}[F] \\ F + 1 \geq 0, \quad \mathcal{H}[F] + 1 \geq 0, \quad \mathcal{C}[F] \geq 0. \quad (13)$$

#### 5.2. Constraints and Energies for Anti-aliasing

By changing the energy functional  $\mathcal{E}$  and constraints  $\mathcal{C}$ , we can tune the properties of  $F(\mathbf{v})$  and thus  $P(\mathbf{v}) = F(\mathbf{v}) + 1$ . We summarize the properties shaped by the constraints and energies considered in Figure 3. These closely follow the criteria analyzed in Section 4, with a low energy low frequency region, bounded global maximum and local maxima.





**Figure 5:** Power spectra generated by minimizing oscillation (top) and total variation (bottom) energies under the low frequency constraint (Section 5.2). Minimizing total variation provides a non-traditional spectrum with a decaying square wave form and lower maximum values for the power spectra.

### Low Frequency Constraint

We start with the property that  $P(\nu)$  has small energy for  $\nu < \nu_0$  for a given  $\nu_0$  (we call this region as the *low frequency region*, and  $\nu > \nu_0$  as the *high frequency region*). This can be imposed with a direct constraint of the form  $\frac{1}{\nu_0} \int_0^{\nu_0} P(\nu) d\nu \leq e_0$ , for a limit  $e_0$ . Similar terms have been used to quantify the energy in the low frequency region in previous works [HSD13, KTBV16a]. However, this does not limit the value of  $P(\nu)$  at a given frequency, and hence  $P(\nu)$  can grow very large, leading to severely high error at certain low frequencies, and hence significant fluctuations in the spectrum. An example spectrum generated with this integral constraint is shown in Figure 4, left. Instead, we propose to directly limit the spectrum for the low frequency region with

$$P(\nu) + 1 \leq e_0 \quad \nu < \nu_0. \quad (14)$$

This ensures that error will be bounded at all low frequencies (Figure 4, right). Ideally,  $e_0 = 0$ , and thus  $P(\nu) = 0$  for  $\nu < \nu_0$ . We will see in the next section that  $e_0$  and hence noise for low frequencies can be traded off with aliasing at higher frequencies. For the analysis in this section, we assume  $e_0 = 0$ .

### Oscillation Energy

Another desired property of  $P(\nu)$  is that it should not have high global and local maxima. This can be imposed in several ways. Previous works [HSD13] have considered measuring squared deviation of  $P(\nu)$  from 1, which can be written as the following energy

$$\mathcal{E}[F] = \int_{\nu_0}^{\infty} F^2(\nu) d\nu. \quad (15)$$

Minimizing this energy with the realizability constraints and the low frequency constraint above for different  $\nu_0$ 's, we get the spectra in Figure 5, top. We get a perfectly zero region for  $\nu < \nu_0$ , and peaks of decaying magnitude for higher frequencies. This is a typically encountered profile for blue noise patterns, except for two recent works [HSD13, KTBV16a]. For  $\nu_0 < \sqrt{1/\pi}$ , which is the theoretical limit for a step-noise profile, we get a perfect step shape. Larger  $\nu_0$  leads to oscillations, with the magnitude of the first peak

determining the global maximum of  $P$ . As  $\nu_0$  is increased, the maximum also gets larger.

### Total Variation Energy

For  $\nu > \nu_0$  and  $\nu_0 > \sqrt{1/\pi}$ ,  $P$  will inevitably deviate from 1 with one or more peaks, before it (possibly) converges to 1 [HSD13]. An alternative way of limiting the magnitudes of these peaks is to minimize total variation energy. This can be visualized as minimizing the length of the path traveled by a point when projected onto the  $P$ -axis, as it moves along the curve  $P(\nu)$  from  $\nu_0$  to  $\infty$ . The resulting energy is given by

$$\mathcal{E}[F] = \int_{\nu_0}^{\infty} |F'(\nu)| d\nu. \quad (16)$$

We show power spectra generated by minimizing this energy under the low frequency constraint and realizability conditions in Figure 5, bottom. The spectra now mostly contain raised rectangular regions instead of peaks, i.e. a decaying square wave. This is due to the sparse gradients introduced by total variation. Heights of rectangular regions, and thus the maxima of  $P$  are smaller than when minimizing the oscillation energy above.

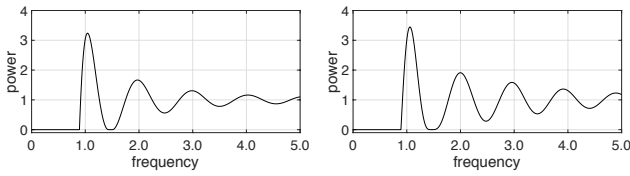
### Smoothness Energy

We further experimented with smoothness energies, the Dirichlet energy  $\mathcal{E}[F] = \int_{\nu_0}^{\infty} |F'(\nu)|^2 d\nu$ , and Laplacian energy  $\mathcal{E}[F] = \int_{\nu_0}^{\infty} |F''(\nu)|^2 d\nu$ . We show the resulting spectra in Figure 6. The spectra in these cases are worse with higher peaks than with oscillation or total variation energy.

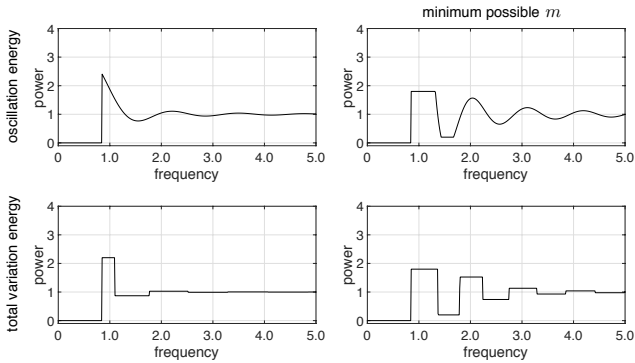
### Maximum Constraint

Although total variation energy leads to peaks of smaller magnitude, it might still be possible to further reduce the global maximum  $m$  of  $P$ . The same is true for all energies. Having a small  $m$  is an important factor to avoid colored noise and hence aliasing, as elaborated on in Section 4.2. To achieve a smaller  $m$ , we limit the magnitude of deviations from 1 with the following constraint

$$|P(\nu) - 1| = |F(\nu)| \leq m_0 - 1 \quad \nu > \nu_0. \quad (17)$$



**Figure 6:** Power spectra generated by minimizing smoothness energies (left: Dirichlet, right: Laplacian energy) for  $\nu_0 = 0.9$ . We get higher peaks and thus worse characteristics than other energies considered.

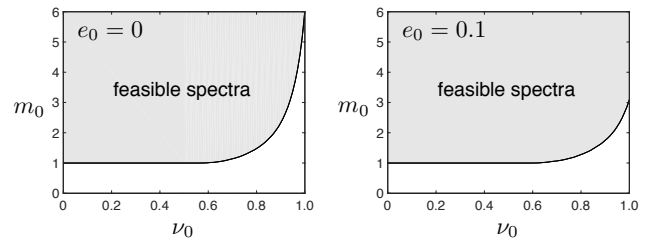


**Figure 7:** Power spectra generated by minimizing oscillation and total variation energies without (left) and with (right) the maximum constraint (Equation 17) for the optimal (minimum possible)  $m_0$  and thus  $m$  ( $\nu_0 = 0.85$  for all cases).

In practice, this constraint is equivalent to  $P(\nu) \leq m_0$ , as for all power spectra in previous works and in this work, the first peak has the largest  $|P(\nu) - 1|$ , and at that peak  $P(\nu) > 1$  (e.g. Figure 5). Of course, not all  $\nu_0 - m_0$  combinations are realizable (we will elaborate more on this point in the next section). In order to define the range of possible  $m_0$  for a given  $\nu_0$ , we can find the minimum possible  $m_0$  by an exhaustive search. We show this constraint imposed on the spectra for oscillation energy and total variation energy in Figure 7. For both energies, reducing  $m_0$  comes at the cost of a larger number of oscillations, albeit all with smaller magnitudes, hence not leading to significant aliasing.

### 5.3. Optimized Sampling Patterns

The analysis above suggests that total variation energy leads to better profiles for power spectra, with lower global and local maxima. Without the maximum constraint (Equation 17), the maximum with oscillation energy is larger (Figure 7, left), while imposing the maximum constraint results in further peaks of higher magnitudes than those with total variation energy as in Figure 7, right (please see the supplementary material for more spectra with total variation and oscillation energies). Due to the flatter shape of the spectrum, noise is introduced for a larger range of unrepresentable high frequencies with total variation energy. However, such incoherent noise is preferred to colored noise caused by higher maxima in power spectra with oscillation energy. Hence, we focus on total variation (Equa-



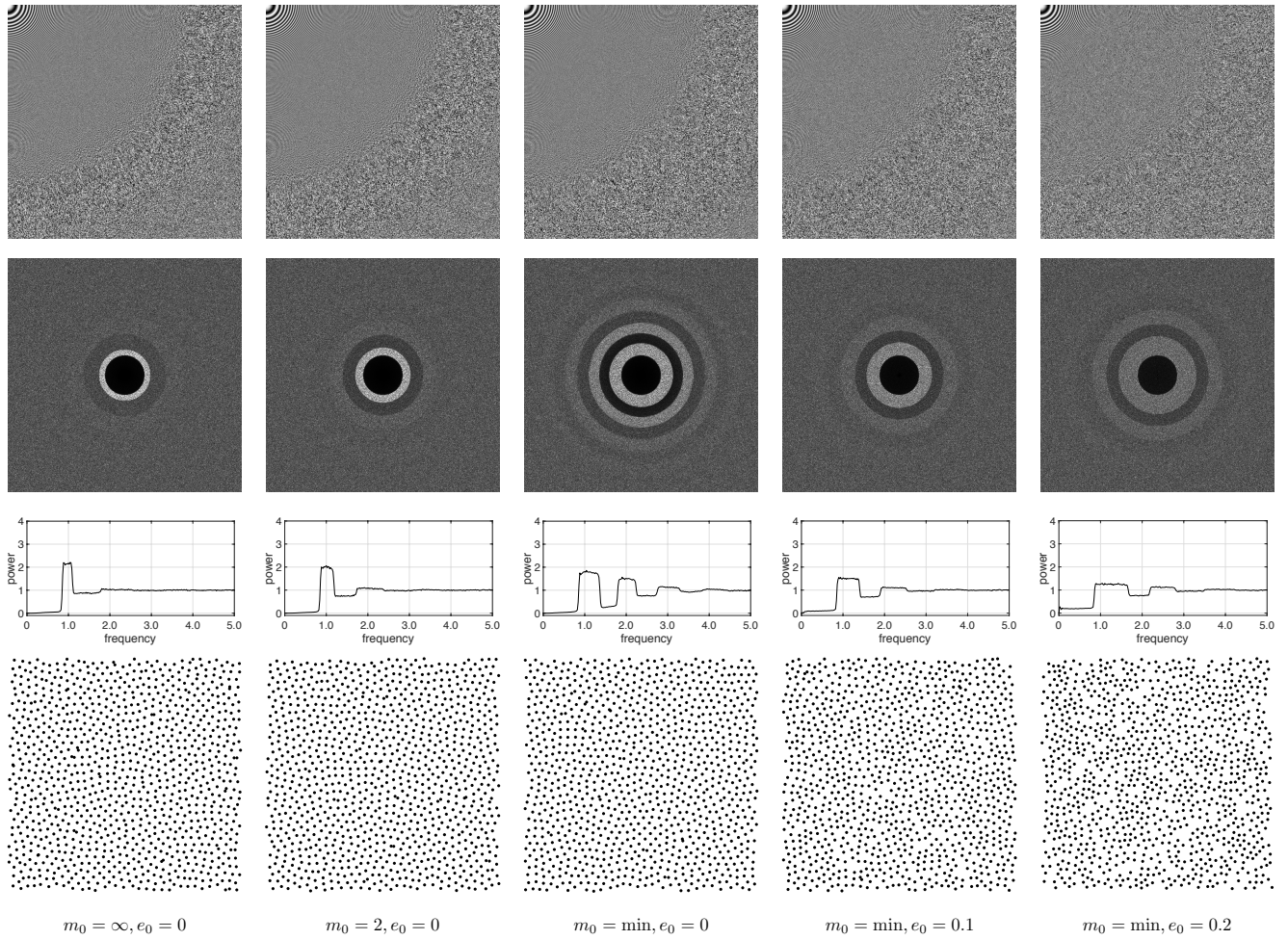
**Figure 8:** Feasible regions for realizable power spectra in the  $\nu_0 - m_0$  space for different  $e_0$ 's. No patterns can have power spectra outside this region.

tion 16) as the energy in this paper. Due to its shape resembling a decaying square wave, we call the resulting pattern as *ds-wave* sampling.

The use of the maximum constraint depends on the gain we obtain, i.e. how much lower the maximum of the power spectrum is with this constraint. We show estimated power spectra without ( $m_0 = \infty$ ) and with the maximum constraint for  $m_0 = 2$ , and  $m_0 = \min$  (the minimum possible  $m_0$ ) in Figure 9. These are computed as the empirical power spectra of generated point distributions (we elaborate more on this in the next section). As we use total variation energy, we already get low maxima, hence see only a marginal improvement. In general, it is possible to tune  $m_0$  depending on how critical this improvement is for the application, but this requires a search among different  $m_0$  values, and some additional local maxima appear in the spectrum.

Recent works [HSD13, KTBV16a] explore minimizing the maximum of power spectra for a given  $\nu_0$  and  $e_0$ . However, as they use pre-defined parametric families of functions, it is not possible to achieve the minimum possible maximum. By utilizing the proposed optimization framework, we can derive the *minimal* maximum (up to numerical accuracy). For fixed  $\nu_0$  and  $e_0$ , we try to optimize any of the above energies for various  $m_0$ 's, and take the minimum that leads to a feasible solution. The resulting space of feasible  $\nu_0 - m_0$  pairs are shown in Figure 8 for  $e_0 = 0$  and  $e_0 = 0.1$ . Note that these results are general and independent of the form of the power spectrum. The minimum possible  $m_0$  stays at 1 for  $\nu < \sqrt{1/\pi}$  as expected, and becomes increasingly more sensitive to  $\nu_0$  for large  $\nu_0$  values. It is not possible to go beyond  $\nu_0 = 1$ , as this is the  $\nu_0$  of regular sampling. We also observe this in practice when we compute the feasible region. We can use the feasible region as a benchmark for how patterns perform for anti-aliasing.

As Figure 8, right, shows, increasing low frequency noise with  $e_0 = 0.1$  significantly reduces  $m_0$ , especially for high values of  $\nu_0$ . We show power spectra obtained for  $e_0 = 0.1$  and  $e_0 = 0.2$  in Figure 9. Although the constraint is  $P(\nu) \leq e_0$  for  $\nu < \nu_0$ , the optimizations result in  $P(\nu) = e_0$  for all  $\nu < \nu_0$ . For higher  $e_0$ 's, power spectra get flatter, and thus zone plate images show reduced aliasing artifacts for high frequencies. This comes at the cost of higher levels of noise introduced into low frequencies, as visible in low frequency parts of the zone plate images. The generated point distributions reveal the source of this low frequency noise: they become increasingly more random for larger values of  $e_0$ .



**Figure 9:** Optimized power spectra with total variation energy and  $v_0 = 0.85$ , for different  $m_0$  and  $e_0$  values. From top to bottom: zone plate image sampled with the generated point distributions, estimated 2D power spectrum, 1D power spectrum, an example distribution.

#### 5.4. Implementation

We discretize the problem in Equation 13 with standard techniques from numerical analysis. As the function  $F$  to optimize for is 1-dimensional, a simple discretization with regular sampling is used. The derivative operators are then discretized with finite differences, and integrals in the energies are approximated with the trapezoidal rule. Hankel transform is discretized with an accurate approximation based on the trapezoidal rule (e.g. [CB93], Equation 6).

We experimented with various ranges and sampling rates for  $F$  and  $\mathcal{H}[F]$ . In all of our experiments, a sample spacing of 0.01 was sufficient for accurate numerical results. In order to test the accuracy of the resulting spectra with this spacing, we take 100 spectra with randomly chosen parameters ( $v_0 \in [0.5, 1]$ , and  $e_0 \in [0, 0.2]$  to avoid having many step-like spectra). We then compute the average root mean squared difference between each spectrum for a sample spacing of 0.01 and 0.001, resulting in  $6.1 \cdot 10^{-3}$  average difference. For  $P$ , we sample the range  $v \in [0, 10]$ , since well before  $v = 10$ ,  $F$  converges to 0 (and hence  $P = F + 1$  converges to 1)

for all patterns in our experiments. The average absolute deviation of  $F$  from 0 for  $v \in [9, 10]$  for 100 random spectra is  $7.74 \cdot 10^{-5}$ . The same, however, is not true for  $\mathcal{H}[F]$ , which determines the PCF  $g$ . For  $\mathcal{H}[F]$ , we thus sample almost the full range of possible distances for the unit toroidal domain we consider,  $[0, 0.5]$  in absolute coordinates. Note, however, that in practice this is not strictly needed as we only require  $g(r) = \mathcal{H}[F](r) + 1 \geq 0$ , and  $g(r)$  does not oscillate significantly beyond a limited range of  $r$ 's.

Hankel transform is a linear operator and hence all constraints, including the realizability conditions (Equation 12), are linear inequality constraints. The oscillation (Equation 15) and smoothness energies turn into quadratic forms when discretized. For these energies, the discrete problem thus becomes quadratic programming. Minimization with total variation energy (Equation 16) can be formulated as linear programming, using well-known results from optimization. Both problems are thus convex and easy to solve with any modern optimization package. We use built-in Matlab functions for optimization.



There are several techniques for synthesis of point patterns based on PCF or power spectrum [ZHW12, OG12, HSD13, WPC\*14, AHD15, KTBV16a]. We experimented with several of these approaches that focus on accuracy [OG12, HSD13, KTBV16a], and got similar results. We hence use the PCF-based matching technique of Heck et al. [HSD13] for all experiments in the paper and the supplementary material, due to the efficient implementation available. We use the default parameters for that algorithm, and the same discretization for PCF as we describe above.

## 6. Evaluation and Analysis

To evaluate the performance of our sampling patterns in practice, we analyze power spectra, and illustrate their anti-aliasing properties on sampled images (for more results, please see the supplementary material).

### Evaluation

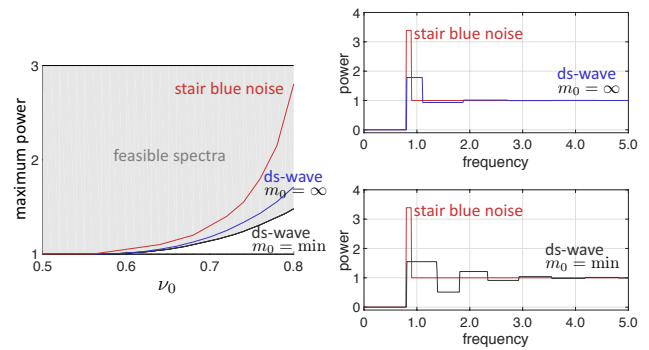
For estimating the power spectrum of a pattern, we generate 10 point distributions with a matching spectrum, each with 4096 points. The empirical power spectra of these distributions, computed with Equation 2 and their radial averages, are averaged to generate all estimated 2D and 1D power spectra in this paper and the supplementary material. For other sampling patterns that start from a theoretical power spectrum (step blue noise, single-peak blue noise [HSD13], and stair blue noise [KTBV16a]), we use the same point distribution synthesis algorithm [HSD13] with the same settings as in Section 5.4.

Unless stated otherwise, we use distributions with 16384 ( $128 \times 128$ ) points at 1 sample per pixel (spp) for all test images except the zone plate images. We use the zone plate function  $(1 + \cos(\alpha\|\mathbf{x}\|^2))/2$ ,  $\mathbf{x} \in \mathcal{V}$  as a benchmark test image, since it reveals aliasing at different frequencies without the masking effect due to local structures [HSD13, KTBV16a]. These images are sampled at 2 spp (please see the supplementary material for zone plate images with 1, 2, 4 spp). All images are reconstructed with a low-pass filter (we use the  $3 \times 3$  Gaussian filter) and resampled to a regular grid. This filter and others used in the literature retain some of the frequencies that cannot be represented with the pixel grid, but ensure a clear visualization of noise. In practice, we only noticed faint and distinct secondary rings for zone plate images when using more than 1 samples per pixel due to this filtering, as visible e.g. in Figure 1.

### Ds-wave and Stair Blue Noise

We analyze the properties of our ds-wave sampling and compare them to those of the state-of-the-art stair blue noise sampling [KTBV16a] in Figure 10. For stair blue noise, at each  $\nu_0$ , we find the minimum value for the maximum  $m$  of the power spectrum with an exhaustive search over the parameters, as described by Kaikhura et al. [KTBV16a]. Since stair blue noise has zero energy for the low frequency region, we set  $e_0 = 0$  for all comparisons.

Figure 10, left shows that for all  $\nu_0 > \sqrt{1/\pi}$  (the theoretical limit for step blue noise), ds-wave sampling has a lower  $m$  than stair blue noise, with the difference getting larger for larger  $\nu_0$ . Ds-wave sampling with  $m_0 = \min$  achieves the optimal  $m$ , by definition. But



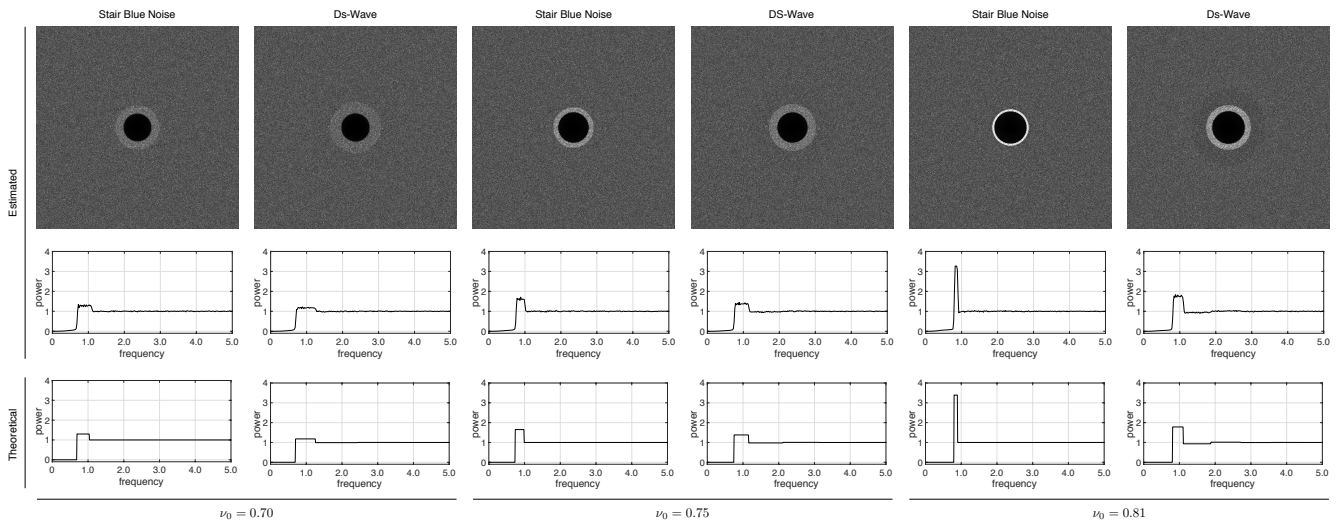
**Figure 10:** (Left) For each  $\nu_0$ , the maximum value of the power spectrum for ds-wave and stair blue noise [KTBV16a] sampling. (Right) Comparisons of spectra of stair blue noise and ds-wave sampling at the maximum achievable  $\nu_0 = 0.81$  for stair blue noise.

even without the maximum constraint (i.e.  $m_0 = \infty$ ), ds-wave sampling has  $m$  very close to the optimum. Note that the maximum  $\nu_0$  we could obtain for stair blue noise sampling for  $m \leq 100$  is 0.81, and thus we show the range  $\nu_0 \in [0.5, 0.8]$ . In Figure 10, right, we plot the spectra for  $\nu_0 = 0.81$ , illustrating the significant difference between ds-wave and stair blue noise for both  $m_0 = \infty$  and  $m_0 = \min$ . We show further theoretical, and estimated 2D and 1D spectra for lower  $\nu_0$ 's in Figure 11 ( $m_0 = \infty$ ). For all cases, the theoretical spectra of ds-wave sampling can be realized reliably, with a lower maximum than stair blue noise, and almost no further oscillations. The difference is especially significant for higher values of  $\nu_0$ . Please see the supplementary material for more examples of theoretical and estimated spectra.

The practical utility of these results is illustrated in Figures 1 ( $\nu_0 = 0.80$ ,  $m_0 = \infty$ ), 2 ( $\nu_0 = 0.80$ ,  $m_0 = \infty$ ), and 12 ( $\nu_0 = 0.85$ ,  $m_0 = \min$ ). In Figure 1, the zone plate image reveals frequencies that are mapped to colored noise and hence secondary patterns due to the high magnitude region for stair blue noise sampling. Ds-wave sampling maps such unrepresentable high frequency content to noise with a profile closer to that for step blue noise, while preserving the same  $\nu_0$ , and thus noise levels for low frequency content, as stair blue noise. Although noise is introduced into a larger range of higher frequencies, this noise is much less objectionable than the patterns introduced by stair blue noise. We illustrate this further in Figures 2 and 12 for images with repeated structures of several frequencies. The cleaner reconstructions of repeated structures of lower frequencies (top rows) provided by stair blue noise sampling come at the expense of mapping repeated structures of frequencies higher than the representable frequency to colored noise. This manifests itself as secondary patterns and higher levels of noise in the sampled images (bottom rows). Ds-wave sampling leads to as white as possible noise for these cases, combining advantages of step and stair blue noise sampling.

### Ds-wave and Other Patterns

We compare ds-wave sampling to further patterns commonly used for anti-aliasing in Figure 13. Dart throwing results in a relatively



**Figure 11:** Estimated and theoretical power spectra for stair blue noise [KTBV16a] and ds-wave sampling for different  $\nu_0$  values ( $e_0 = 0$ ,  $m_0 = \infty$ ). For all cases, ds-wave sampling results in a lower maximum for the power spectra.

small  $m$  and hence does not lead to objectionable aliasing artifacts. However, it also leads to low frequency noise in sampled images, as is apparent for the zone plate image. For a wider range of cleaner low frequencies for sampled images (i.e. a larger low frequency region in power spectrum), CCCVT centroids [BSD09] can be utilized. This results, however, in higher peak values, and thus more pronounced aliasing artifacts.

An even larger low frequency region is possible with single-peak blue noise sampling [HSD13], as illustrated in Figure 13, middle. Note that single-peak blue noise is not exactly zero at low frequencies due to the introduced Gaussian at around the transition from low to high frequencies, while ds-wave sampling has zero energy in the low frequency region. We set  $\nu_0$  such that both single-peak and ds-wave reach 1 the first time at approximately the same  $\nu$ . At this size of the low frequency region, the peak has a high value and aliasing becomes apparent as secondary patterns in the zone plate image in Figure 13, middle, and the sampled high frequency repeated stripe patterns (bottom rows) in Figure 12. Our ds-wave sampling at  $\nu_0 = 0.85$  and  $e_0 = 0$  has the same size of the low frequency region, but with a lower maximum, and hence leads to lower noise and aliasing, as visible in the same figures.

The maximum  $m$  and hence aliasing artifacts can be further reduced by introducing low frequency noise. With  $e_0 = 0.1$ , we get a slightly smaller  $m$  than dart throwing, while still having cleaner and a larger range of low frequencies than dart throwing as shown in Figure 13.

### Multiple Samples per Pixel

One way of reducing noise is increasing the number of samples per pixel. However, if  $P$  contains high peaks, for finite spp, there will always be secondary patterns due to aliasing when sampling image content of certain frequencies. To see this, we start by noting that the error after resampling to a regular grid is given

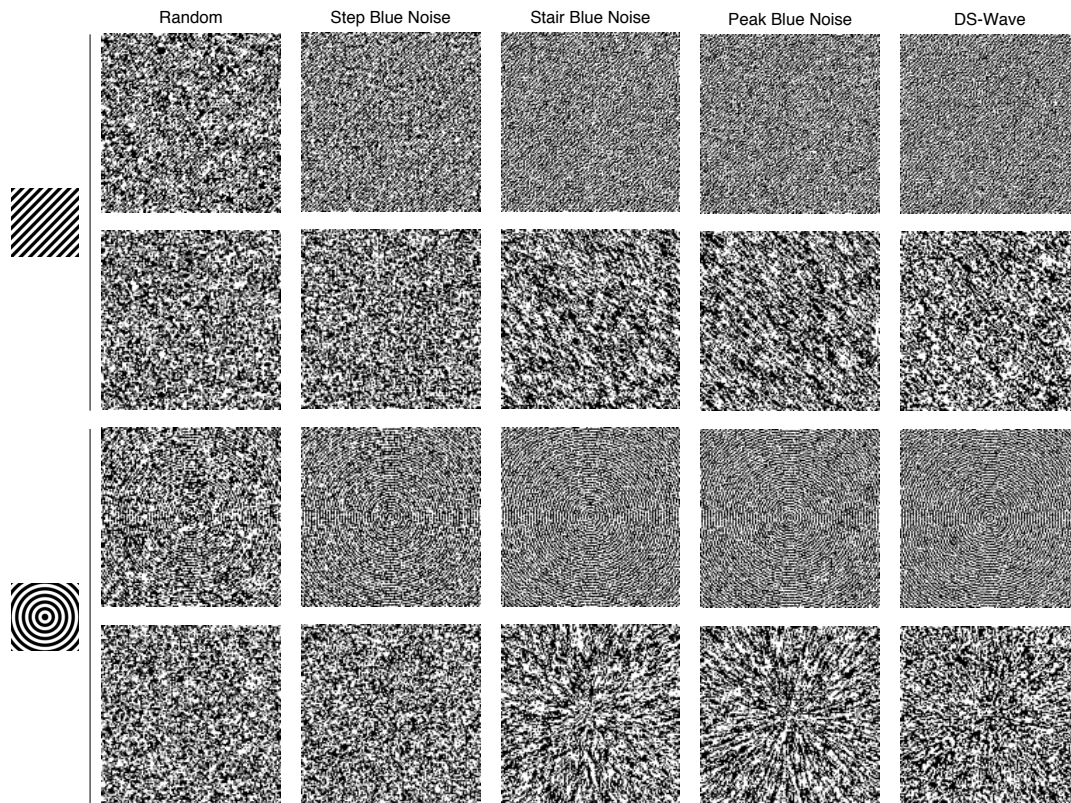
by  $|K(\mathbf{v})|^2 E(\mathbf{v}) = |K(\mathbf{v})|^2 \frac{1}{\lambda} [P_t * (U + 1)](\mathbf{v})$ , as derived in Section 4.1. Increasing spp means we are keeping  $K$  the same, and expanding  $P$  (as it is related to  $g$  with a Fourier transform, which compresses for larger number of points due to smaller distances among them, please see Section 3). If  $P$  has peaks, they will thus be shifted to higher frequencies and be smoothed as a result of this expansion. If a sampled image has local structures of those frequencies, due to the convolution in the definition of  $E(\mathbf{v})$ , these peaks will then be shifted to lower frequencies that are captured by the filter  $K$ . Hence, similar but smoothed artifacts in the form of visible secondary patterns will appear in the final reconstructed image.

This is illustrated in Figure 14 for the cosine function in Figure 2 with  $\nu_c = 0.85$ . We use 4 spp for the top row, and 16 spp for the bottom row. Note that as we always normalize frequencies by the number of sampling points, the absolute frequency shifts with the spp. For both cases, increasing spp does not help to reduce the visible secondary structures due to aliasing. In fact, higher spp might lead to perceptually more apparent secondary structures, e.g. for 16 spp in Figure 14.

### Artifacts on Rendered Images

We illustrate such aliasing artifacts for practical rendered scenes in Figure 15. For this figure, we sample each of the dimensions for the light transport, except the image plane, densely. Hence, the spp reported corresponds to the image plane samples. Computing each of those image plane samples is thus a costly operation involving a numerical integration for all other dimensions. We show a reference image first, then the result of stair blue noise on a smaller image with 1 or 2 spp on the left, and the corresponding result with ds-wave sampling on the right. Note that we intentionally did not re-sample the rendered small images with nearest neighbor sampling to illustrate that applying standard filters on the images with aliasing artifacts does not alleviate the aliasing artifacts due to colored noise. For certain scenes such as the top image, increasing the spp





**Figure 12:** State-of-the-art sampling patterns such as stair blue noise [KTBV16a] and peak blue noise [HSD13] result in less noise than random and step blue noise sampling for repeated structures with a low frequency (top rows), at the cost of introducing colored noise and hence secondary patterns when sampling images with repeated structures of higher frequencies (bottom rows). Ds-wave sampling leads to cleaner reconstructions, and smoothly degrades to as white as possible noise for repeated structures with high frequencies, avoiding visible secondary patterns. (Visualizations of the original stripe patterns in sampled images are shown on the left. The patterns repeat every  $1/v_c$  pixels, with  $v_c = 0.5$  (top rows) and  $v_c = 0.9$  (bottom rows). Peak blue noise is with an effective  $v_0 = 0.85$  as explained in the text, stair blue noise is with  $v_0 = 0.81$  (maximum possible), and ds-wave is with  $v_0 = 0.85$ ,  $m_0 = \min$ .)

from 1 to 2 makes the aliasing artifacts more apparent. In general, we observed that ds-wave sampling makes the most difference for directional repeated structures as exemplified in the figure.

### Running Time

The formulation of the optimization problems with linear and quadratic programming allows us to use efficient and robust solvers. For total variation energy (linear programming), it takes 1-5 minutes for the solver to converge on a PC with Intel(R) Xeon(R) CPU ES-2680 v3 @ 2.5 GHz, with the running time increasing for larger  $v_0$ . As this optimization is done once and offline, the main computational complexity comes from the point distribution generation procedure [HSD13], which takes about one minute to converge for 4096 points.

## 7. Conclusions and Future Work

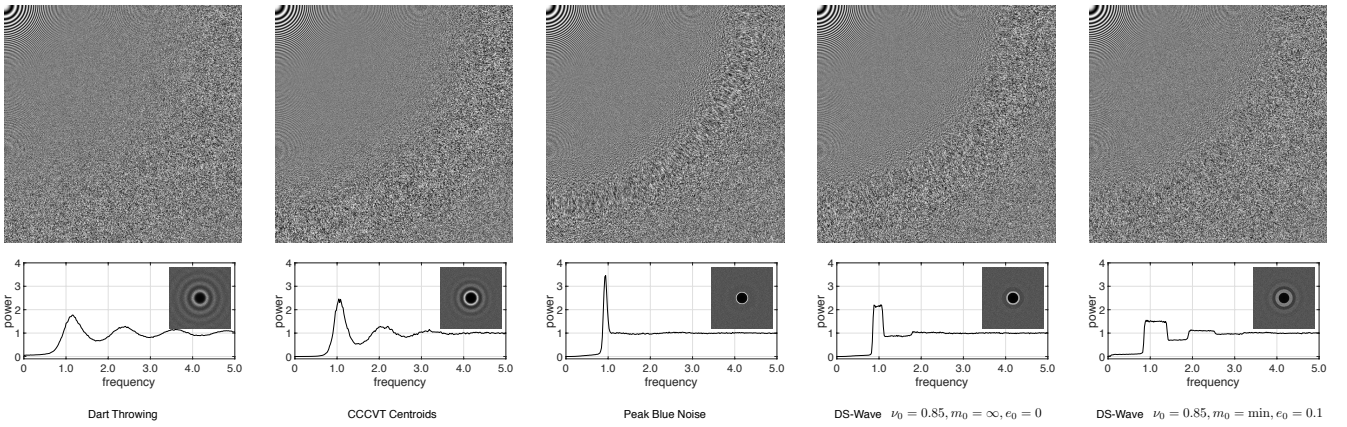
We presented a theoretical and practical framework for analyzing aliasing, and generating sampling patterns with optimized properties for anti-aliasing via formulating the problem of generating re-

alizable spectra as variational optimization. The resulting patterns lead to practical improvements in reducing aliasing artifacts due to colored noise, and the proposed theoretical framework allows us to explore and revise optimality measures used for anti-aliasing. We see many interesting uses of this framework for future research, some of which we summarize below.

**Sampling for Integration** Although we focused on anti-aliasing when reconstructing images in the scope of this paper, a very promising direction is to optimize power spectra for reducing error in numerical integration. Recent works [Ö16, PSC\*15] have proved that the dependence of total error in integration on power spectrum is given by  $\frac{1}{\lambda} \int_{-\infty}^{\infty} P_t(\mathbf{v})(U(\mathbf{v}) + 1)d\mathbf{v}$  (which is in contrast with the error *spectrum* for representation in Equation 8 as we also discussed in Section 4.1). Minimizing this error will turn into a linear programming problem when formulated as variational optimization, similar to Equation 13.

**Adaptive Anti-aliasing Patterns** Similar to previous works, we considered non-adaptive anti-aliasing patterns, with no information on the actual image to be represented. Recent works [ROG17]





**Figure 13:** Reconstructed zone plate images and estimated power spectra for different sampling patterns. Ds-wave sampling can get the same effective range of the low frequency region as single-peak blue noise [HSD13] with significantly less aliasing artifacts, which can be reduced even further by introducing low frequency noise with  $e_0 = 0.1$ .

show that sampling patterns with adaptive second order product densities can lead to significant accuracy improvements for image representation and processing. By combining our optimization framework with locally adaptive point distribution synthesis algorithms [ROG17], we can obtain optimal adaptive sampling patterns for image reconstruction.

**Exploration of Second Order Characteristics** Previous works explore the space of valid second order characteristics either via analysis of available point patterns [OG12], or parameterized families of power spectra [HSD13, KTBV16a]. Our framework can be used to explore this space without such constraints. As an example, we showed that optimal maximal values for power spectra for given  $\nu_0$ 's can be obtained (Figure 8). Similar results can be derived for other applications such as geometry sampling, physically-based simulations, or natural distributions.

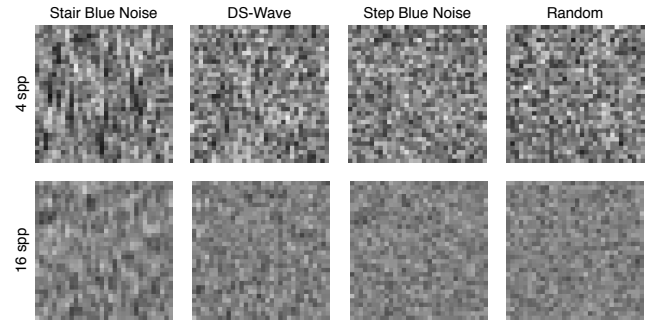
**Higher Dimensional Sampling** An interesting aspect of the optimization problem in Equation 13 is that it depends on the dimensionality due to the Hankel transform. We will thus get *different spectra for different dimensions*, as the Hankel transform takes a different form for different dimensions. It will be interesting to explore optimal sampling patterns for higher dimensions, e.g. in the context of rendering where the integrands can be very high dimensional.

## Appendix A: Derivation of Error Spectrum

We start by rewriting the form of the error in Equation 7

$$E = \frac{1}{\lambda^2} \mathbb{E}_{\mathcal{P}} [ |S * T|^2 ] + |T|^2 - \frac{2}{\lambda} \Re \{ (\mathbb{E}_{\mathcal{P}} [S] * T) \bar{T} \}. \quad (18)$$

As defined in Section 3,  $s(\mathbf{x}) = \sum_j \delta(\mathbf{x} - \mathbf{x}_j)$  and thus its Fourier transform is  $S(\mathbf{v}) = \sum_j e^{-2\pi i \mathbf{v}^T \mathbf{x}_j}$ . Plugging this into the Campbell's theorem of first order (Equation 3 in Appendix A of the supplementary material) we get



**Figure 14:** Increasing number of samples per pixel does not fundamentally solve the aliasing problem. We illustrate this when sampling the cosine function in Figure 2 with  $\nu_c = 0.85$  with stair blue noise sampling ( $\nu_0 = 0.80$ ) for 4 and 16 spp. In contrast, ds-wave ( $\nu_0 = 0.80$ ,  $m_0 = \infty$ ), step blue noise, and random sampling do not lead to secondary patterns but to incoherent noise.

tary material) we get

$$\mathbb{E}_{\mathcal{P}} [S(\mathbf{v})] = \mathbb{E}_{\mathcal{P}} \left[ \sum_j e^{-2\pi i \mathbf{v}^T \mathbf{x}_j} \right] = \lambda \int_{\mathcal{V}} e^{-2\pi i \mathbf{v}^T \mathbf{x}} d\mathbf{x} = \lambda \delta(\mathbf{v}). \quad (19)$$

The last term in Equation 18 thus becomes  $-2[\delta * T](\mathbf{v}) \overline{T(\mathbf{v})} = -2|T(\mathbf{v})|^2$ . Calculating the first term in Equation 18 is more involved due to the squared magnitude. We first expand this term with

the definition of  $S$  and utilizing properties of the Fourier transform

$$\begin{aligned}
|[S * T](\mathbf{v})|^2 &= \left| \sum_j e^{-2\pi i \mathbf{v}^T \mathbf{x}_j} * T(\mathbf{v}) \right|^2 \\
&= \sum_{jk} \overline{\left( e^{-2\pi i \mathbf{v}^T \mathbf{x}_j} * T(\mathbf{v}) \right)} \left( e^{-2\pi i \mathbf{v}^T \mathbf{x}_k} * T(\mathbf{v}) \right) \\
&= \sum_{jk} t(\mathbf{x}_j) t(\mathbf{x}_k) e^{-2\pi i \mathbf{v}^T (\mathbf{x}_k - \mathbf{x}_j)} \\
&= \sum_j t^2(\mathbf{x}_j) + \sum_{j \neq k} t(\mathbf{x}_j) t(\mathbf{x}_k) e^{-2\pi i \mathbf{v}^T (\mathbf{x}_k - \mathbf{x}_j)},
\end{aligned} \tag{20}$$

where we used the notation  $a(\mathbf{v}) * b(\mathbf{v})$  for  $[a * b](\mathbf{v})$ , and the equivalence  $e^{-2\pi i \mathbf{v}^T \mathbf{x}_j} * T(\mathbf{v}) = \mathcal{F}[\delta(\mathbf{x} - \mathbf{x}_j)t(\mathbf{x})] = \mathcal{F}[\delta(\mathbf{x} - \mathbf{x}_j)t(\mathbf{x}_j)] = e^{-2\pi i \mathbf{v}^T \mathbf{x}_j} t(\mathbf{x}_j)$ . The expected value of the first term on the last line can be computed with Campbell's theorem of first order as  $\mathbb{E}_{\mathcal{P}} \left[ \sum_j t^2(\mathbf{x}_j) \right] = \lambda \int_{\mathcal{V}} t^2(\mathbf{x}) d\mathbf{x}$ . The second term involves a double sum, and the expected value can thus be computed by utilizing Campbell's theorem of second order (Equation 4 of Appendix A in the supplementary material) as

$$\begin{aligned}
\mathbb{E}_{\mathcal{P}} \left[ \sum_{j \neq k} t(\mathbf{x}_j) t(\mathbf{x}_k) e^{-2\pi i \mathbf{v}^T (\mathbf{x}_k - \mathbf{x}_j)} \right] \\
&= \lambda^2 \int_{\mathcal{V} \times \mathcal{V}} t(\mathbf{x}) t(\mathbf{y}) e^{-2\pi i \mathbf{v}^T (\mathbf{x} - \mathbf{y})} g(\mathbf{x} - \mathbf{y}) d\mathbf{x} d\mathbf{y} \\
&= \lambda^2 \int_{\mathcal{V}} a_t(\mathbf{r}) e^{-2\pi i \mathbf{v}^T \mathbf{r}} g(\mathbf{r}) d\mathbf{r} \\
&= \lambda^2 \mathcal{F}[a_t g](\mathbf{v}) = \lambda^2 \left[ |T|^2 * G \right](\mathbf{v}),
\end{aligned} \tag{21}$$

with  $a_t(\mathbf{r})$  denoting the autocorrelation of  $t(\mathbf{x})$ , and we use the relation  $\mathcal{F}[a_t](\mathbf{v}) = |T(\mathbf{v})|^2$ , and the multiplication theorem of Fourier transform. Substituting the expression for  $G$  (Equation 4), this can also be written in terms of  $U$  as  $\lambda^2 \left[ |T|^2 * (U/\lambda + \delta) \right](\mathbf{v}) = \lambda \left[ |T|^2 * U \right](\mathbf{v}) + \lambda^2 |T(\mathbf{v})|^2$ . Summing the two terms in Equation 20, we thus get

$$\mathbb{E}_{\mathcal{P}} \left[ |[S * T](\mathbf{v})|^2 \right] = \lambda \int_{\mathcal{V}} t^2(\mathbf{x}) d\mathbf{x} + \lambda \left[ |T|^2 * U \right](\mathbf{v}) + \lambda^2 |T(\mathbf{v})|^2. \tag{22}$$

Finally, we sum all the terms in Equation 18

$$\begin{aligned}
E(\mathbf{v}) &= \frac{1}{\lambda^2} \left( \lambda \int_{\mathcal{V}} t^2(\mathbf{x}) d\mathbf{x} + \lambda \left[ |T|^2 * U \right](\mathbf{v}) + \lambda^2 |T(\mathbf{v})|^2 \right) \\
&\quad + |T(\mathbf{v})|^2 - 2|T(\mathbf{v})|^2 \\
&= \frac{1}{\lambda} \left( \int_{\mathcal{V}} t^2(\mathbf{x}) d\mathbf{x} + \left[ |T|^2 * U \right](\mathbf{v}) \right) \\
&= \frac{1}{\lambda} \left( \left[ |T|^2 * \mathbf{1} \right](\mathbf{v}) + \left[ |T|^2 * U \right](\mathbf{v}) \right) \\
&= \frac{1}{\lambda} \left[ |T|^2 * (U + \mathbf{1}) \right](\mathbf{v}).
\end{aligned} \tag{23}$$

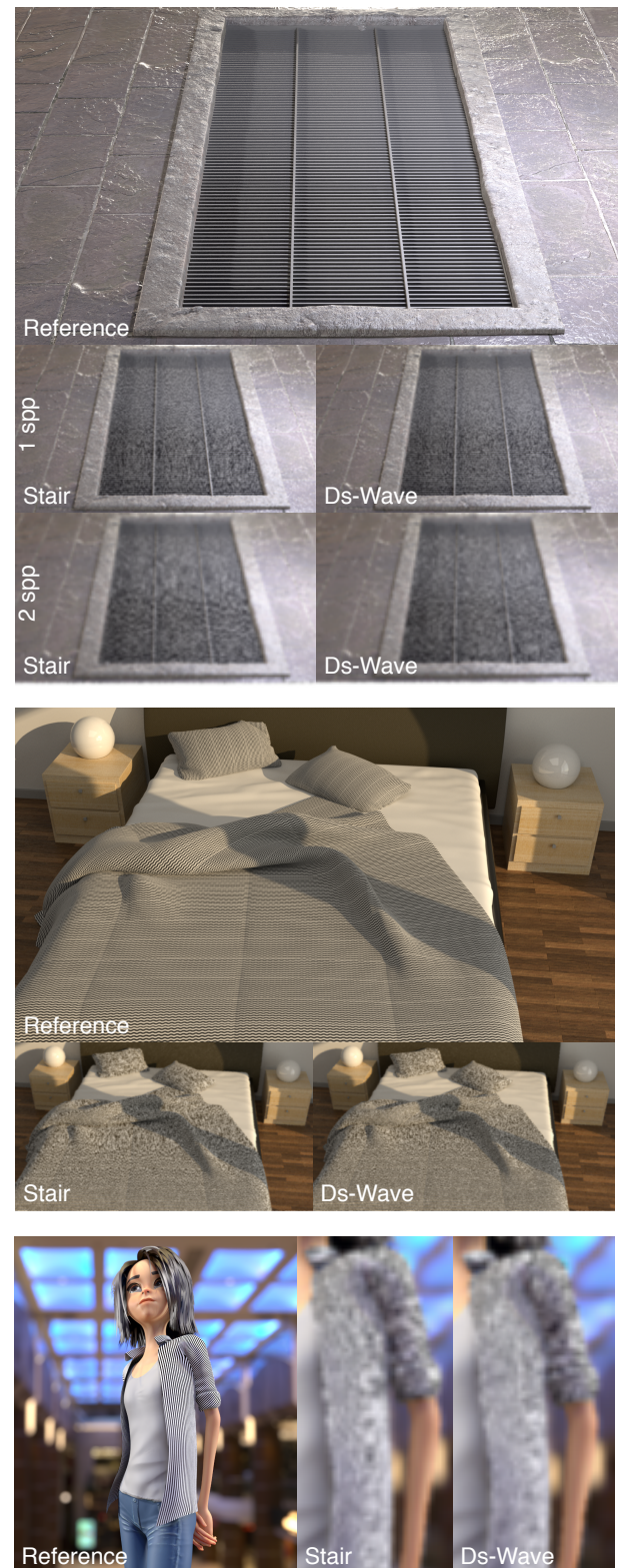
## References

[AHD15] AHMED A. G. M., HUANG H., DEUSSEN O.: Aa patterns for point sets with controlled spectral properties. *ACM Trans. Graph.* 34, 6 (Oct. 2015), 212:1–212:8. 2, 3, 11

- [Bri07] BRIDSON R.: Fast poisson disk sampling in arbitrary dimensions. In *ACM SIGGRAPH 2007 Sketches* (New York, NY, USA, 2007), SIGGRAPH '07, ACM. 2
- [BSD09] BALZER M., SCHLÖMER T., DEUSSEN O.: Capacity-constrained point distributions: A variant of lloyd's method. *ACM Trans. Graph.* 28, 3 (July 2009), 86:1–86:8. 2, 12
- [CB93] CREE M., BONES P.: Algorithms to numerically evaluate the hankel transform. *Computers & Mathematics with Applications* 26, 1 (1993), 1–12. 10
- [Coo86] COOK R. L.: Stochastic sampling in computer graphics. *ACM Trans. Graph.* 5, 1 (1986), 51–72. 2
- [CYC\*12] CHEN Z., YUAN Z., CHOI Y.-K., LIU L., WANG W.: Variational blue noise sampling. *IEEE Transactions on Visualization and Computer Graphics* 18, 10 (Oct. 2012), 1784–1796. 2
- [dGBOD12] DE GOES F., BREEDEN K., OSTROMOUKHOV V., DESBRUN M.: Blue noise through optimal transport. *ACM Trans. Graph.* 31, 6 (Nov. 2012), 171:1–171:11. 2
- [DH06] DUNBAR D., HUMPHREYS G.: A spatial data structure for fast poisson-disk sample generation. *ACM Trans. Graph.* 25 (July 2006), 503–508. 2
- [Dur11] DURAND F.: *A frequency analysis of Monte-Carlo and other numerical integration schemes*. Tech. Rep. MIT-CSAILTR-2011-052, CSAIL, MIT, MA, February 2011. 4
- [DW85] DIPPÉ M. A. Z., WOLD E. H.: Antialiasing through stochastic sampling. *SIGGRAPH Comput. Graph.* 19, 3 (July 1985), 69–78. 2, 3, 4, 7
- [EMP\*12] EBEIDA M. S., MITCHELL S. A., PATNEY A., DAVIDSON A. A., OWENS J. D.: A simple algorithm for maximal poisson-disk sampling in high dimensions. *Comput. Graph. Forum* 31, 2pt4 (2012), 785–794. 2
- [EPM\*14] EBEIDA M. S., PATNEY A., MITCHELL S. A., DALBEY K. R., DAVIDSON A. A., OWENS J. D.: K-d darts: Sampling by k-dimensional flat searches. *ACM Trans. Graph.* 33, 1 (Feb. 2014), 3:1–3:16. 2
- [Fat11] FATTAL R.: Blue-noise point sampling using kernel density model. *ACM Trans. Graph.* 30, 4 (July 2011), 48:1–48:12. 2
- [HSD13] HECK D., SCHLÖMER T., DEUSSEN O.: Blue noise sampling with controlled aliasing. *ACM Trans. Graph.* 32, 3 (July 2013), 25:1–25:12. 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14
- [IPSS08] ILLIAN J., PENTTINEN A., STOYAN H., STOYAN D. (Eds.): *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley and Sons, Ltd., 2008. 2, 3
- [JZW\*15] JIANG M., ZHOU Y., WANG R., SOUTHERN R., ZHANG J. J.: Blue noise sampling using an sph-based method. *ACM Trans. Graph.* 34, 6 (Oct. 2015), 211:1–211:11. 2
- [KCODL06] KOPF J., COHEN-OR D., DEUSSEN O., LISCHINSKI D.: Recursive wang tiles for real-time blue noise. *ACM Trans. Graph.* 25, 3 (July 2006), 509–518. 2
- [KTBV16a] KAILKHURA B., THIAGARAJAN J. J., BREMER P.-T., VARSHNEY P. K.: Stair blue noise sampling. *ACM Trans. Graph.* 35, 6 (Nov. 2016), 248:1–248:10. 1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14
- [KTBV16b] KAILKHURA B., THIAGARAJAN J. J., BREMER P. T., VARSHNEY P. K.: Theoretical guarantees for poisson disk sampling using pair correlation function. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (March 2016), pp. 2589–2593. 2
- [LD08] LAGAE A., DUTRÉ P.: A comparison of methods for generating Poisson disk distributions. *Comput. Graph. Forum* 27, 1 (March 2008), 114–129. 2
- [MEA\*18] MITCHELL S. A., EBEIDA M. S., AWAD M. A., PARK C., PATNEY A., RUSHDI A. A., SWILER L. P., MANOCHA D., WEI L.-Y.: Spoke-darts for high-dimensional blue-noise sampling. *ACM Trans. Graph.* 37, 2 (May 2018). 2



- [Mit87] MITCHELL D. P.: Generating antialiased images at low sampling densities. *SIGGRAPH Comput. Graph.* 21, 4 (Aug. 1987), 65–72. 2
- [Mit91] MITCHELL D. P.: Spectrally optimal sampling for distribution ray tracing. *SIGGRAPH Comput. Graph.* 25, 4 (July 1991), 157–164. 2, 6
- [MW04] MØLLER J., WAAGEPETERSEN R. P.: *Statistical inference and simulation for spatial point processes*. Chapman & Hall/CRC, 2003, Boca Raton (Fl.), London, New York, 2004. 3
- [Ö16] ÖZTIRELI A. C.: Integration with stochastic point processes. *ACM Trans. Graph.* 35, 5 (Aug. 2016), 160:1–160:16. 4, 6, 13
- [OG12] ÖZTIRELI A. C., GROSS M.: Analysis and synthesis of point distributions based on pair correlation. *ACM Trans. Graph.* 31, 6 (Nov. 2012), 170:1–170:10. 2, 3, 11, 14
- [Ost07] OSTROMOUKHOV V.: Sampling with polyominoes. *ACM Trans. Graph.* 26, 3 (July 2007). 2
- [PSC\*15] PILLEBOUE A., SINGH G., COEURJOLLY D., KAZHDAN M., OSTROMOUKHOV V.: Variance analysis for monte carlo integration. *ACM Trans. Graph.* 34, 4 (July 2015), 124:1–124:14. 4, 6, 13
- [ROG17] ROVERI R., ÖZTIRELI A. C., GROSS M.: General Point Sampling with Adaptive Density and Correlations. *Computer Graphics Forum* (2017). 13, 14
- [SGBW10] SCHMALTZ C., GWOSDEK P., BRUHN A., WEICKERT J.: Electrostatic halftoning. *Comput. Graph. Forum* 29, 8 (2010), 2313–2327. 2
- [SHD11] SCHLÖMER T., HECK D., DEUSSEN O.: Farthest-point optimized point sets with maximized minimum distance. In *Proceedings of the ACM SIGGRAPH Symposium on High Performance Graphics* (New York, NY, USA, 2011), HPG '11, ACM, pp. 135–142. 2
- [SK13] SUBR K., KAUTZ J.: Fourier analysis of stochastic sampling strategies for assessing bias and variance in integration. *ACM Trans. Graph.* 32, 4 (July 2013), 128:1–128:12. 4
- [TS02] TORQUATO S., STILLINGER F. H.: Controlling the short-range order and packing densities of many-particle systems. *The Journal of Physical Chemistry B* 106, 43 (2002), 11406–11406. 4
- [Uli88] ULICHNEY R.: Dithering with blue noise. *Proceedings of the IEEE* 76, 1 (Jan 1988), 56–79. 2
- [UST06] UCHE O., STILLINGER F., TORQUATO S.: On the realizability of pair correlation functions. *Physica A: Statistical Mechanics and its Applications* 360, 1 (2006), 21–36. 2, 3
- [Wei08] WEI L.-Y.: Parallel poisson disk sampling. *ACM Trans. Graph.* 27 (August 2008), 20:1–20:9. 2
- [Wei10] WEI L.-Y.: Multi-class blue noise sampling. *ACM Trans. Graph.* 29, 4 (July 2010), 79:1–79:8. 2
- [WPC\*14] WACHTEL F., PILLEBOUE A., COEURJOLLY D., BREEDEN K., SINGH G., CATHELIN G., DE GOES F., DESBRUN M., OSTROMOUKHOV V.: Fast tile-based adaptive sampling with user-specified fourier spectra. *ACM Trans. Graph.* 33, 4 (July 2014), 56:1–56:11. 2, 3, 11
- [WW11] WEI L.-Y., WANG R.: Differential domain analysis for non-uniform sampling. *ACM Trans. Graph.* 30, 4 (July 2011), 50:1–50:10. 2
- [XLGG11] XU Y., LIU L., GOTSMAN C., GORTLER S. J.: Capacity-constrained delaunay triangulation for point distributions. *Computers & Graphics* 35, 3 (2011), 510–516. Shape Modeling International (SMI) Conference 2011. 2
- [Yel83] YELLOTT J.: Spectral consequences of photoreceptor sampling in the rhesus retina. *Science* 221, 4608 (1983), 382–385. 2
- [Yuk15] YUKSEL C.: Sample elimination for generating poisson disk sample sets. *Computer Graphics Forum* 34, 2 (2015), 25–32. 2
- [ZHWW12] ZHOU Y., HUANG H., WEI L.-Y., WANG R.: Point sampling with general noise spectrum. *ACM Trans. Graph.* 31, 4 (July 2012), 76:1–76:11. 2, 3, 11



**Figure 15:** For each image, a reference rendering, and rendered images with 1 or 2 spp with stair blue noise, and ds-wave sampling are shown.