Supplementary Document for Practical Coding Function Design for Time-of-Flight Imaging

Felipe Gutierrez-Barragan, Syed Azer Reza, Andreas Velten, Mohit Gupta
University of Wisconsin-Madison
{fgutierrez3, sreza2, velten}@wisc.edu, mohitg@cs.wisc.edu

S. 1. Quantifying the Performance of a CW-ToF Coding Scheme

In this supplementary section, we provide a detailed description and derivation of the mean expected absolute depth error of a CW-ToF coding scheme.

A general metric to evaluate the performance of a coding scheme is the *mean expected absolute depth error*, first introduced in [3]. Consider the 3-D space consisting of all possible unknown vectors $\mathbf{U} = [d, I_s, I_a]$ (distance d, effective albedo I_a, and ambient brightness I_a). Each coding function pair, $M_i(t)$ and $D_i(t)$, will map some $\mathbf{U}$ to an intensity measurement I_i. A decoding algorithm, $\mathcal{D}$, is used to map the vector of K brightness measurements $\mathcal{D} : \mathbf{I} \rightarrow \mathbf{U}$. Due to different noise sources such as shot noise, sensor read noise, quantization, and saturation, the sensor measures a noisy intensity $\hat{\mathbf{I}} = \mathbf{I} + \eta$. The decoding algorithm will then perform a noisy mapping of $\hat{\mathbf{I}}$ to an estimated vector of unknowns $\hat{\mathbf{U}}$. Since $\hat{\mathbf{I}}$ is a random variable we can calculate the *expected absolute depth error* for a given unknown vector $\mathbf{U}$ and coding scheme as follows:

$$\mathbb{E}[|d - \hat{d}|] = \int_{\hat{\mathbf{I}}} |d - \hat{d}| p(\hat{\mathbf{I}}|\mathbf{I}) \mathrm{d}\hat{\mathbf{I}} \tag{1}$$

where $p(\hat{\mathbf{I}}|\mathbf{I})$ is the probability density function of the measured noisy brightness $\hat{\mathbf{I}}$, given the true brightness $\mathbf{I}$. Equation 1 gives the expected depth error for a fixed unknown $\mathbf{U}$ and coding scheme. To evaluate the performance of a coding scheme we can calculate the *mean expected absolute depth error* over the full depth range,

$$\boxed{\frac{1}{d_{max}} \int_0^{d_{max}} \mathbb{E}[|d - \hat{d}|] \mathrm{d}d = \frac{1}{d_{max}} \int_0^{d_{max}} \int_{\hat{\mathbf{I}}} |d - \hat{d}| p(\hat{\mathbf{I}}|\mathbf{I}) \mathrm{d}\hat{\mathbf{I}} \mathrm{d}d} \tag{2}$$

Equation 2 is the mean expected absolute depth error for a fixed effective albedo I_s and ambient component I_a. For simplicity, in this paper we refer to the mean expected absolute depth error as the mean depth error. Each point in the mean depth error surface plots shown in the paper and in this report (e.g. Figures 4, 6, 7) correspond to an evaluation of Equation 2 for a given I_s and I_a. Note that the total source power and the ambient illumination form part of I_s and I_a, respectively. Consequently, we only need to vary those two terms to generate the mean depth error surface plots.

Computation of the expected absolute depth error: Since we know $p(\hat{\mathbf{I}}|\mathbf{I})$ we can efficiently compute the first integral in Equation 2 via Monte Carlo integration. This is done by sampling $\hat{\mathbf{I}}$ from $p(\hat{\mathbf{I}}|\mathbf{I})$, computing $\hat{d}$, and approximating $\mathbb{E}[|d - \hat{d}|]$ using those samples. For our simulations we found that $20,000$ samples sufficed to give a stable $\mathbb{E}[|d - \hat{d}|]$.

Normalized cross-correlation for depth decoding: The above description of the mean depth error computation assumes that the decoding algorithm $\mathcal{D}$ is known. A naive computation of $\hat{d}$ given $\hat{\mathbf{I}}$ for an arbitrary coding scheme would involve a computationally expensive 3-dimensional search over the space of all possible unknowns $\mathbf{U}$. Fortunately, we can greatly reduce the computational complexity of depth decoding by following a similar near-optimal decoding approach as in [6]. Let $\mathbf{C}$ be the $N \times K$ matrix where each column is a correlation function χ_i for a given coding scheme. The phase shift and consequently $\hat{d}$ of the modulation function will be proportional to the argument that maximizes the zero-mean normalized cross correlation (ZNCC) between $\mathbf{C}$ and $\hat{\mathbf{I}}$, as shown in Equation 3. This procedure is analog to the standard ZNCC template matching algorithm commonly used in image processing [9].

$$\hat{d} \sim \underset{1 \leq i \leq N}{\operatorname{argmax}} \frac{\mathbf{C}_i - \operatorname{mean}(\mathbf{C}_i)}{\|\mathbf{C}_i - \operatorname{mean}(\mathbf{C}_i)\|} \cdot \frac{\hat{\mathbf{I}} - \operatorname{mean}(\hat{\mathbf{I}})}{\|\hat{\mathbf{I}} - \operatorname{mean}(\hat{\mathbf{I}})\|} \tag{3}$$

The above decoding algorithm works for arbitrary coding schemes and has linear time complexity ($\mathcal{O}(N)$). For certain coding schemes, such as sinusoid coding, the decoding algorithm has an analytical expression, which effectively reduces the run-time complexity to $\mathcal{O}(1)$. For any practical implementation of a coding scheme, the run-time complexity of their decoding algorithms is one additional trade-off that needs to be considered. Previous coding function design work has demonstrated that efficient depth decoding algorithms for a specific coding scheme can be developed through the geometric analysis of the correlation functions [3] or through regression methods [1].

S. 2. How does the Performance of Sinusoid Coding Change with the Number of Measurements?

In this section we show that if the total integration time is fixed and split evenly among K intensity measurements the performance of a sinusoid coding scheme remains constant as K varies. We follow a similar analysis as in [3].

We want to *bound* the expected absolute depth error, $\mathbb{E}[|d-\hat{d}|]$, due to intensity measurement error $\mathbb{E}[|\mathbf{I}-\hat{\mathbf{I}}|]$. Furthermore, we want to show that as we vary K the bound does not change for a sinusoid coding scheme. We start by considering the first-order taylor series approximation of an intensity measurement about the point $d = \hat{d}$:

$$\mathbf{I}(d) = \mathbf{I}(\hat{d}) + \frac{\partial \mathbf{I}}{\partial d}(\hat{d})(d - \hat{d}) \tag{4}$$

where $\frac{\partial \mathbf{I}}{\partial d}(\hat{d}) = I_s \chi'(\hat{d})$. For simplicity of the derivation we assume that the effective albedo I_s does not depend on d [1]. Let $\mathbf{I}(\hat{d}) = \hat{\mathbf{I}}$ and $\mathbf{I}(d) = \mathbf{I}$. We can re-arrange Equation 4 and take the magnitude of both sides in the following way:

$$|d - \hat{d}| = \frac{|\mathbf{I} - \hat{\mathbf{I}}|}{|I_s \chi'(\hat{d})|} = \frac{\sqrt{\sum_{i=1}^{K}(I_i - \hat{I}_i)^2}}{I_s \sqrt{\sum_{i=1}^{K}(\chi_i'(\hat{d}))^2}} \tag{5}$$

Equation 5 gives us a relationship between the absolute depth error ($|d - \hat{d}|$) due to small measurement noise ($|\mathbf{I} - \hat{\mathbf{I}}|$). Since both $\hat{\mathbf{I}}$ and $\hat{d}$ are random variables we can take the expectation of both sides of the expression,

$$\mathbb{E}[|d - \hat{d}|] = \mathbb{E}[\frac{\sqrt{\sum_{i=1}^{K}(I_i - \hat{I}_i)^2}}{I_s \sqrt{\sum_{i=1}^{K}(\chi_i'(\hat{d}))^2}}] \approx \frac{\mathbb{E}[\sqrt{\sum_{i=1}^{K}(I_i - \hat{I}_i)^2}]}{\mathbb{E}[I_s \sqrt{\sum_{i=1}^{K}(\chi_i'(\hat{d}))^2}]} \tag{6}$$

where the above equality comes from the first-order approximation of the expectation of the ratio of two random variables [8, 2, 10]. Recall from our Gaussian noise model that $\hat{I}_i \sim N(I_i, I_i + \sigma_{read}^2)$ [4], where the variance terms I_i and σ_{read}^2 correspond are associated to shot noise and sensor read noise respectively. Therefore we can calculate bound the expectation of the numerator and calculate it in the following way:

$$\mathbb{E}[\sqrt{\sum_{i=1}^{K}(I_i - \hat{I}_i)^2}] \leq \sqrt{\sum_{i=1}^{K}\mathbb{E}[(I_i - \hat{I}_i)^2]} = \sqrt{\sum_{i=1}^{K}\text{Var}[\hat{I}_i]} = \sqrt{K\sigma_{read}^2 + \sum_{i=1}^{K} I_i} \tag{7}$$

where the above inequality follows from Jensen's inequality on concave function (i.e. square root) and from the linearity of expectation. We can continue bounding the numerator by analyzing the upper-bound of I_i and recalling that $0 \leq D(t) \leq 1$,

$$I_i = I_s \int_0^{\tau} D(t)\bar{M}(t - \frac{2d}{c})dt + I_a \leq I_s \int_0^{\tau} \bar{M}(t - \frac{2d}{c})dt + AT = I_s + AT \tag{8}$$

where the above inequality is the result from upper bounding $D(t) = 1$ and consequently $I_a = A\int_0^T D(t)dt = AT$. Substituting Equation 8 into Equation 7 we get:

$$\mathbb{E}[\sqrt{\sum_{i=1}^{K}(I_i - \hat{I}_i)^2}] \leq \sqrt{K(\sigma_{read}^2 + I_s + AT)} \tag{9}$$

Note that we can only sum up the I_s and I_a terms from each measurement as we did above because we assume that the exposure/integration time T from all measurements is the same. We can now analyze the denominator from Equation 6:

$$\mathbb{E}[I_s \sqrt{\sum_{i=1}^{K}(\chi_i'(\hat{d}))^2}] = I_s(\frac{1}{d_{max}}) \int_0^{d_{max}} \sqrt{\sum_{i=1}^{K}(\chi_i'(\hat{d}))^2} d\hat{d} = \frac{I_s \mathcal{L}_{\text{curve}}}{d_{max}} \tag{10}$$

where we assume that with no prior information the probability of all depths $\hat{d}$ is the same, i.e. $p(\hat{d}) = \frac{1}{d_{max}}$. We also introduce the coding curve length term $\mathcal{L}_{\text{curve}}$ as defined in [3]. Furthermore, we denote d_{max} as the maximum unambiguous

[1] In practice, I_s does depend on scene distances d due to light intensity fall-off

depth determined by the fundamental frequency of the coding scheme. Bringing together the results from Equations 6, 9 and 10 we can define the upper bound on the expected depth error due to small measurement noise:

$$\boxed{\mathbb{E}[|d - \hat{d}|] \leq \frac{d_{max}\sqrt{K(\sigma_{read}^2 + I_s + AT)}}{I_s \mathcal{L}_{\text{curve}}}} \tag{11}$$

Equation 11 is the upper bound of the expected absolute depth error for an arbitrary coding scheme where the total integration time is split evenly among measurements. For a sinusoid coding scheme we can give an explicit bound by plugin in the the length of the sinusoid coding curve $\mathcal{L}_{\text{sinusoid}} = \frac{\sqrt{K}}{4\sqrt{2}}$,

$$\mathbb{E}[|d - \hat{d}|] \leq \frac{(4\sqrt{2})d_{max}\sqrt{\sigma_{read}^2 + I_s + AT}}{I_s} \tag{12}$$

Note that the dependence on K disappears. Hence, the performance of a sinusoid coding scheme does not increase as K increases if the total integration time is fixed and evenly split. This is an important result that adds one additional heuristic in the search of families of high-performance coding functions. The coding curve length of a coding function family should increase faster than $\sqrt{K}$ as K increases. The Hamiltonian coding curve length increase exponentially with K [3]. This means that if we had no hardware constraints, a Hamiltonian coding scheme with a large K could achieve an increasingly small depth error, without having to increase total power and total exposure time.

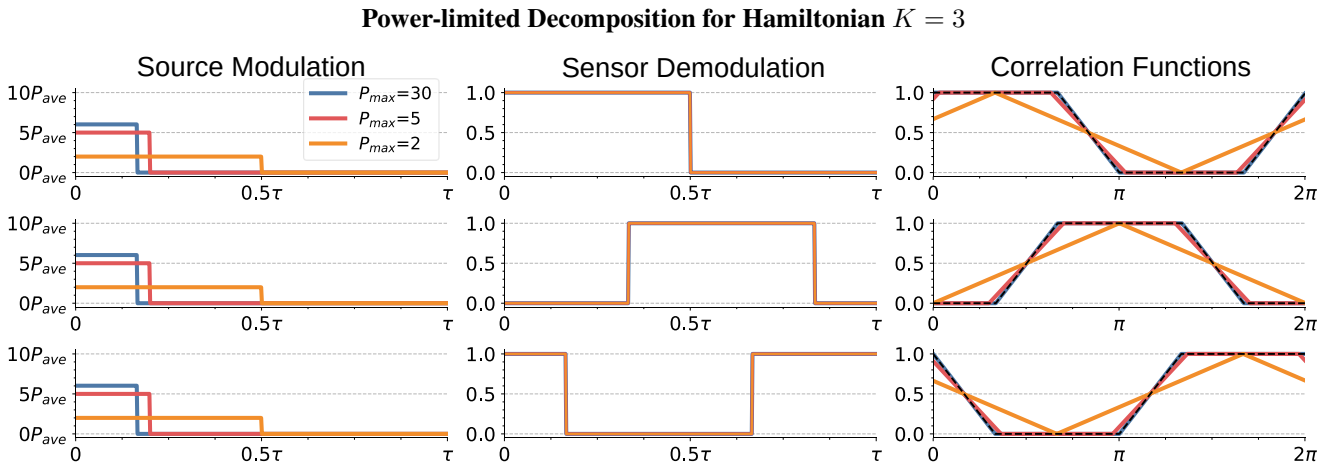
S. 3. Power-limited High-Performance Modulation and Demodulation Functions

In this section we present the correlation decomposition for Hamiltonian $K = [3, 4, 5]$ with a finite light source peak power and infinite bandwidth. Please refer to section 5 of the main document for a detailed discussion of how the results in this section are obtained.

Main observations: Figures 1, 2, and 3 show the power-limited correlation decompositions for Hamiltonian $K = 3$, $K = 4$, and $K = 5$ under three maximum power constraints: $P_{max} = 30P_{ave}$, $P_{max} = 5P_{ave}$, and $P_{max} = 2P_{ave}$. We find that all coding schemes are able to achieve a zero-error decomposition as long as the maximum power constraint is high enough; $6P_{ave}$ for Hamiltonian $K = 3$, $12P_{ave}$ for Hamiltonian $K = 4$, $30P_{ave}$ for Hamiltonian $K = 5$. Interestingly, at maximum power levels where none of the coding schemes achieve a zero error decompositon, we can still find good approximations. For instance, at $P_{max} = 5P_{ave}$, despite the maximum power being 6 times lower than the zero-error threshold we are still able to find good approximations of all correlation functions.

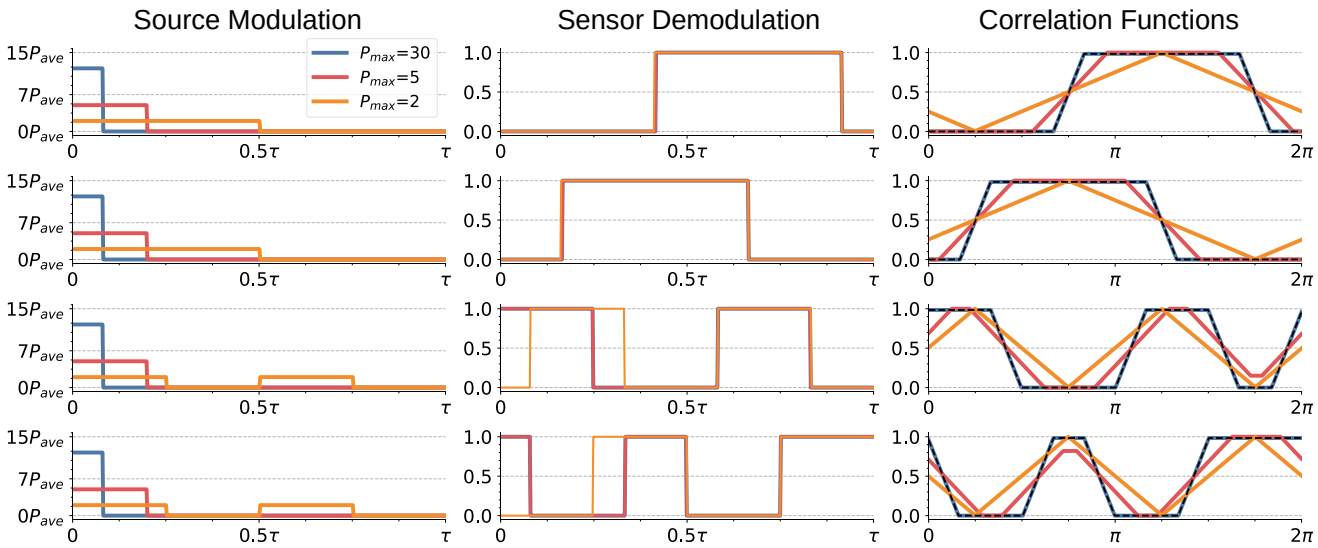
- **Shape of the functions:** All solutions have a binary structure. In particular, a common pattern is shared by all the non-zero error solutions, i.e. solutions that result in approximations of Hamiltonian codes. For a given maximum power constraint, the modulation function will be a binary function with a peak power equal to P_{max}.

Summary: In summary, the resulting power-limited binary high-performance codes could be easily implemented in current CW-ToF cameras. Their binary structure allows them to leverage the digital square wave generators used in CW-ToF cameras. Furthermore, even with strong maximum power constraints (e.g. $P_{max} = 5P_{ave}$) our algorithm finds good approximations to the objective high-performance codes.



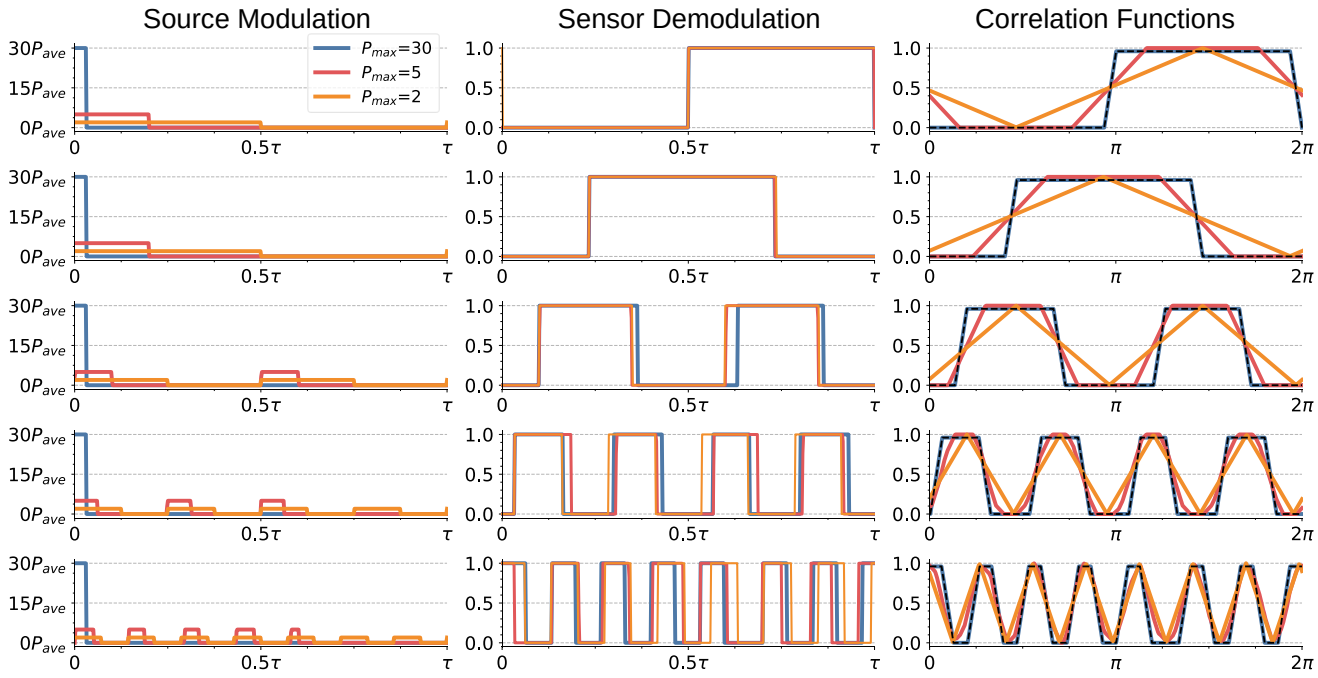
Supplementary Figure 1: **Power-limited decomposition for Hamiltonian $K = 5$.** Power-limited correlation function decomposition for Hamiltonian $K = 3$ at different maximum power levels. At $P_{max} \geq 6P_{ave}$ the resulting modulation and demodulation functions have the same correlation as the objective Hamiltonian correlation. This is clearly illustrated in the case of $P_{max} = 30P_{ave}$ (blue) where the chosen solution only uses $6P_{ave}$. When $P_{max} \leq 6P_{ave}$ only approximations of the objective correlations are obtained.

Power-limited Decomposition for Hamiltonian $K = 4$



Supplementary Figure 2: **Power-limited decomposition for Hamiltonian $K = 4$.** Power-limited correlation function decomposition for Hamiltonian $K = 4$ at different maximum power levels. At $P_{max} \geq 12P_{ave}$ the resulting modulation and demodulation functions have the same correlation as the objective Hamiltonian correlation. This is clearly illustrated in the case of $P_{max} = 30P_{ave}$ (blue) where the chosen solution only uses $12P_{ave}$. When $P_{max} \leq 12P_{ave}$ only approximations of the objective correlations are obtained.

Power-limited Decomposition for Hamiltonian $K = 5$



Supplementary Figure 3: **Power-limited decomposition for Hamiltonian $K = 5$.** Power-limited correlation function decomposition for Hamiltonian $K = 5$ at different maximum power levels. At $P_{max} \geq 30P_{ave}$ the resulting modulation and demodulation functions have the same correlation as the objective Hamiltonian correlation. When $P_{max} \leq 30P_{ave}$ only approximations of the objective correlations are obtained.

S. 4. Performance of Power-limited Coding Functions

In this section we present the performance (mean depth errors) for power-limited Hamiltonian $K = [3, 4, 5]$, sinusoid, and multi-frequency sinusoid.

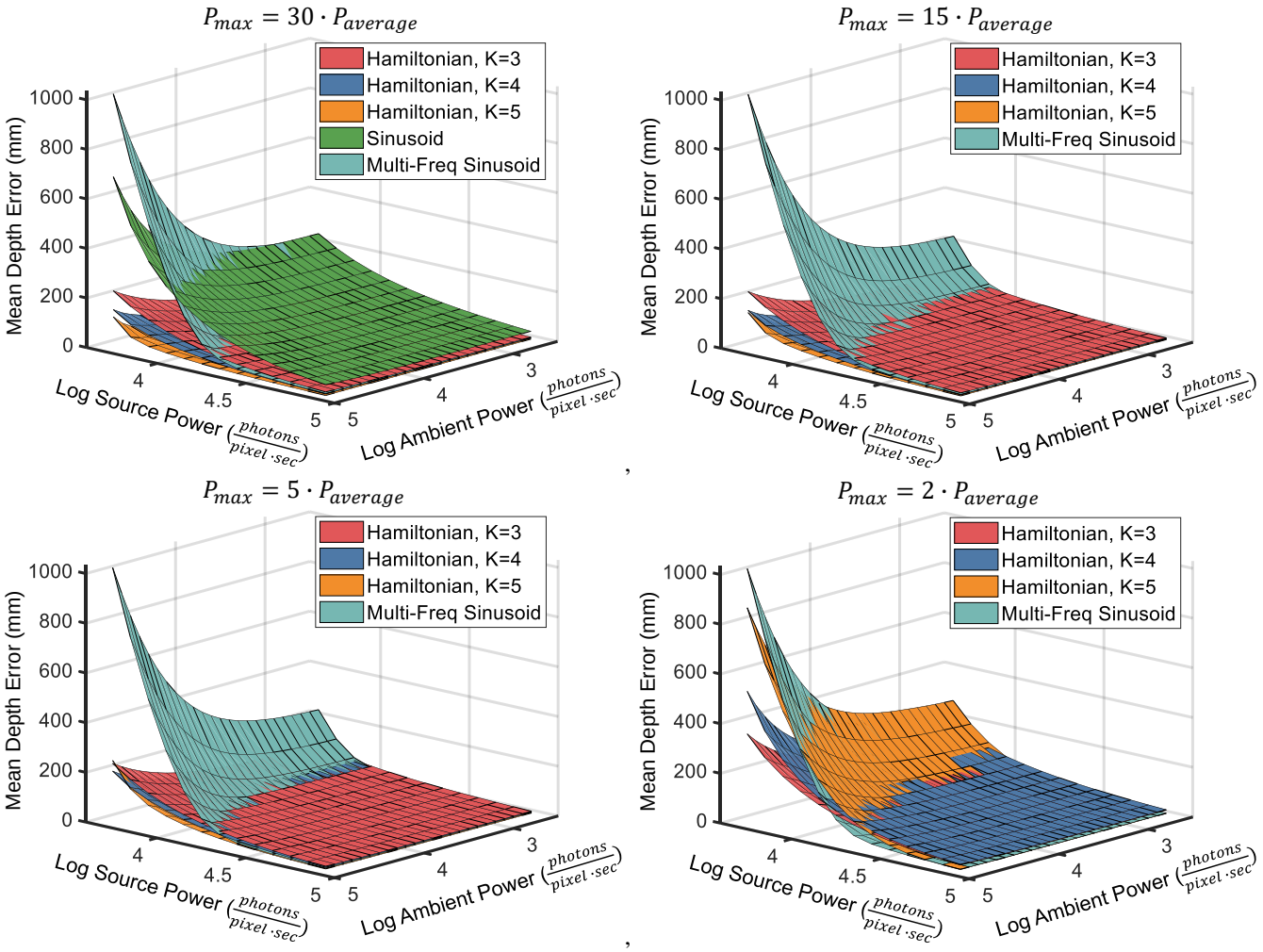
For the simulations we used the following parameters. Fundamental frequency $f_0 = 15$Mhz, corresponding to a 10m depth range. Total integration time of 300ms, and scene albedos of $0 - 1$. The source and ambient average power are varied to emulate a wider range of realistic SNR levels. At each source and ambient power level the mean depth error, as described in Section S. 1, is computed.

Main observations: Figure 4 shows the mean depth error results for power-limited coding schemes at maximum power levels of $30P_{ave}$, $15P_{ave}$, $5P_{ave}$, and $2P_{ave}$. Power-limited Hamiltonian coding functions are robust to low SNR levels, even at strict maximum power levels (e.g. $P_{max} = 5P_{ave}$). On the other hand, the performance of multi-frequency sinusoid rapidly decays as we lower SNR (i.e. lower source power and increase ambient power). In the limiting case of $P_{max} = 2P_{ave}$ the best performing coding scheme depends on the SNR level. At high SNR multi-frequency sinusoid is the best choice and at low SNR Hamiltonian $K = 3$ is the better choice. Note that in that case Hamiltonian $K = 3$ becomes equivalent to square coding (see Figure 1).

- **Crossing points due to phase unwrapping errors:** As we decrease SNR the performance of multi-frequency sinusoid quickly degrades due to phase unwrapping errors. This can be seen from the crossing points in all the mean depth error plots as source power and ambient power are varied. The proposed practical coding schemes are robust to phase unwrapping errors, despite the high-frequency coding functions present in some of them (e.g. Hamiltonian $K = 5$, see the fifth rows of Figure 3).

- **Crossing points due to stronger power constraints:** In the extreme power-limited case where $P_{max} = 2P_{ave}$ there are multiple crossing points across coding schemes as we vary the SNR level. Hamiltonian $K = 3$, however, shows the most stable performance across all SNR levels.

Summary: Overall the proposed power-limited coding schemes outperform sinusoid-based codes at most SNR levels. As long as the maximum power constraint is greater than $2P_{ave}$ the choice of coding scheme is relatively clear (i.e. Hamiltonian $K = 5$).



Supplementary Figure 4: **Mean depth error results for power-limited coding schemes.** Mean depth error simulations for the power-limited coding schemes at maximum power levels of $30P_{ave}$, $15P_{ave}$, $5P_{ave}$, and $2P_{ave}$ and infinite bandwidth. For maximum power levels above $5P_{ave}$ Hamiltonian $K = 5$ outperforms all coding schemes at all SNR levels. In the limiting case when $P_{max} = 2P_{ave}$ Hamiltonian $K = 3$ is the best choice for low SNR and multi-frequency sinusoid is the best choice for high SNR levels.

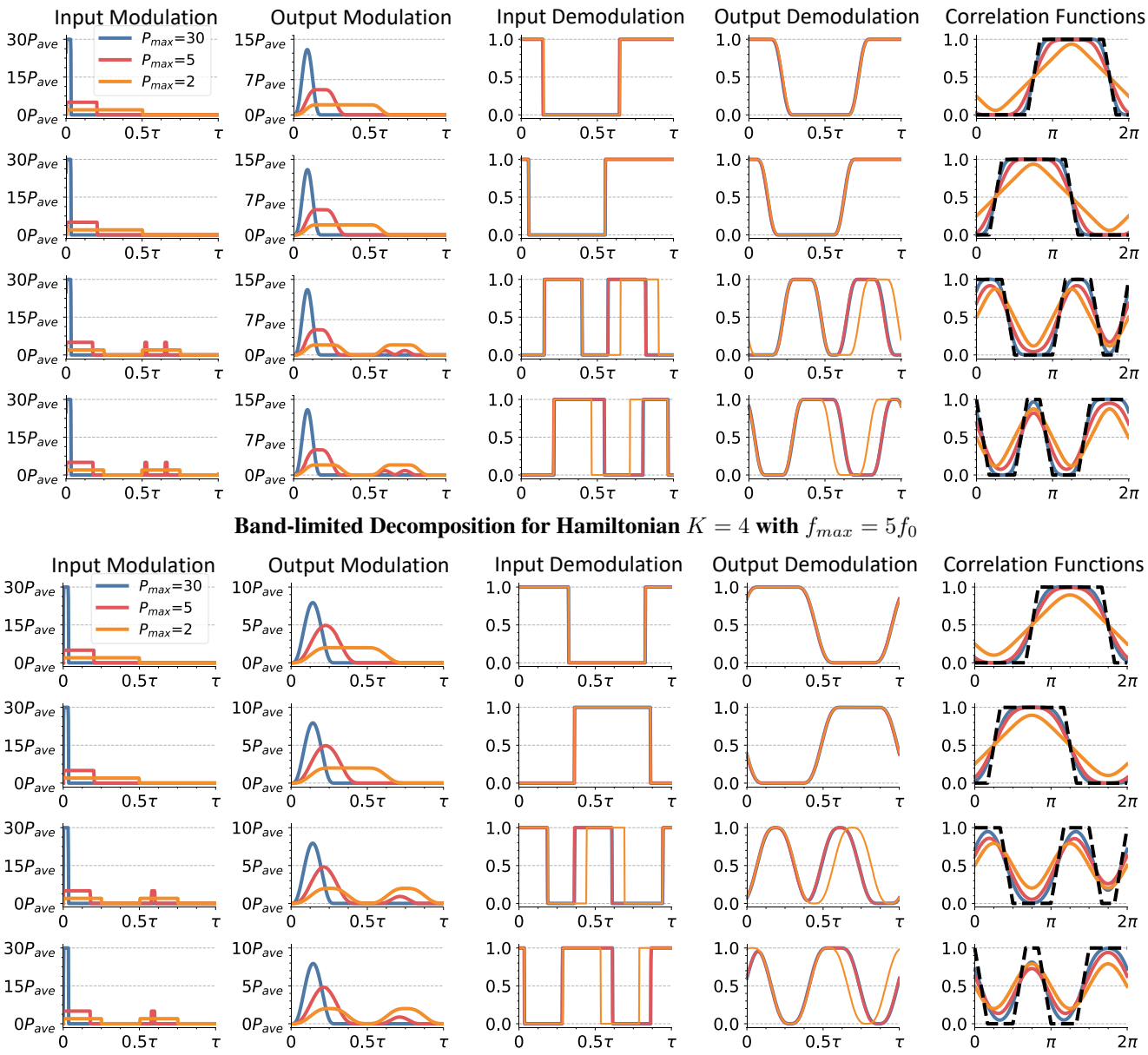
S. 5. Band-limited High-Performance Modulation and Demodulation Functions

In this section we present the correlation decomposition for Hamiltonian $K = [3, 4, 5]$ with a finite light source peak power and finite bandwidth. The binary and maximum power constraints are applied to the *input* modulation and demodulation functions of the CW-ToF system. The bandwidth constraint is applied to the *output* functions which get correlated.

Main observations: In this paper we analyzed the band-limited correlation decomposition for maximum frequencies of: $f_{max} = 10f_0$ and $5f_0$. For each bandwidth constraint we performed the decomposition at maximum power levels of $P_{max} = 30P_{ave}, 5P_{ave}, 2P_{ave}$. Figure 5 shows the band-limited decomposition of Hamiltonian $K = 4$. At $P_{max} = 30P_{ave}$ the bandwidth constraint also imposes an implicit peak power constraint as shown in the transition of the input to the output modulation functions. Therefore, the improvement in the correlation function estimation obtained from increasing the maximum power constraint is minimal, specially in the case of $5f_0$ (e.g. difference between blue and read lines). Similar band-limited decomposition results are obtained for Hamiltonian $K = 3$ and $K = 5$. Please refer to the appendix of this document for these decompositions.

- **Shape of the functions:** Similar to the power-limited results, all solutions for the input modulation and demodulation functions have a binary structure. Hamiltonian coding function require infinite bandwidth to be able to implement them perfectly. Therefore, in a band-limited setting all solutions will be approximations. Interestingly, the same pattern found in the power-limited approximations is present in the band-limited approximations. For a given maximum power constraint, the input modulation function will be a binary function with a peak power equal to P_{max}.

Summary: In summary, despite the fact that Hamiltonian codes require infinite bandwidth to be able to be implemented perfectly, we can achieve good approximations even at strong bandwidth constraints.



Supplementary Figure 5: **Band-limited decomposition for Hamiltonian $K = 4$.**

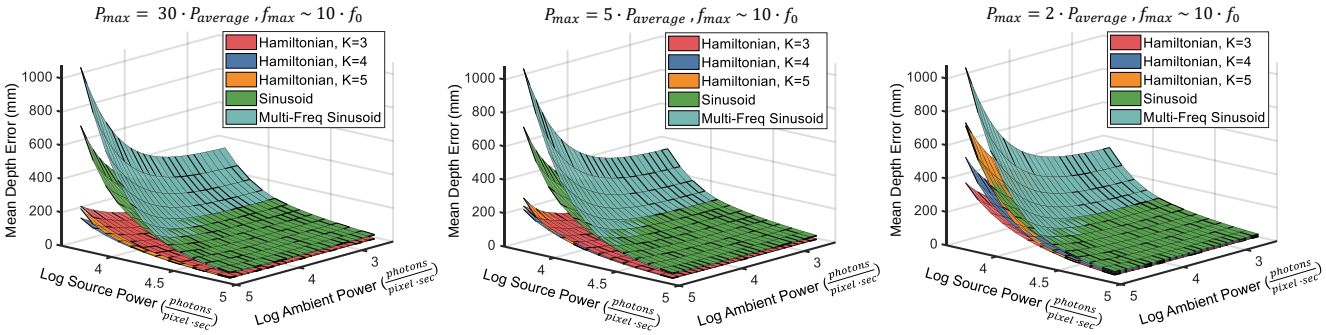
S. 6. Performance of Band-limited Coding Functions

In this section we present the performance (mean depth errors) for band-limited Hamiltonian $K = [3, 4, 5]$, sinusoid, and multi-frequency sinusoid.

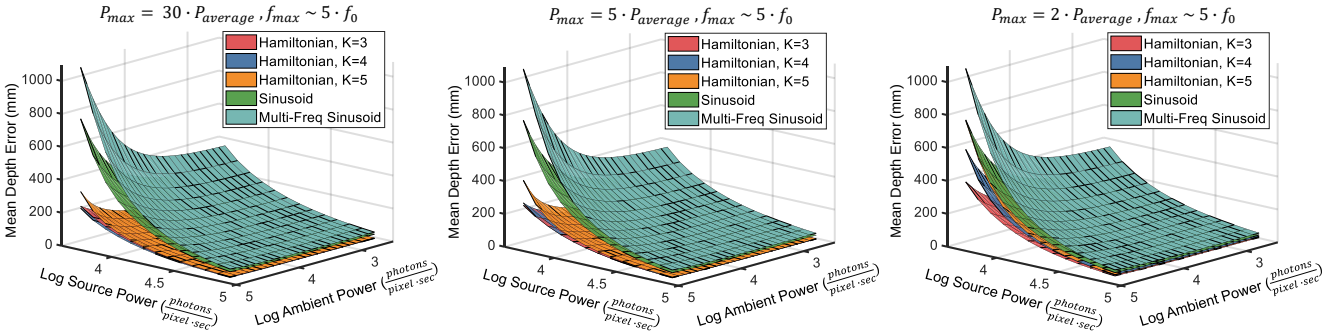
Similar mean depth error simulation parameters are used as in Section S. 4.

Main observations: The performance of all coding schemes is degraded due to the system's bandwidth. Nonetheless, the severity of the degrade in performance varies across coding schemes (e.g. the effect on sinusoid coding is almost negligible). Figures 6 and 7 show that the choice of coding scheme is highly dependent on the hardware constraints. A weak bandwidth constraint (e.g. infinite or large bandwidth) leads to our practical Hamiltonian $K = 5$ being the best choice. As we introduce stronger bandwidth constraints as in Figures 6 and 7, the band-limited Hamiltonian $K = 4$ shows the best performance. However, if the maximum power constraint is lowered too much (e.g. $2P_{ave}$, right-most plot in Figures 6 and 7) the band-limited Hamiltonian $K = 3$ has the best performance. Therefore, in practice, a clear understanding of the system's bandwidth and maximum power constraints is required to choose the coding scheme with the best performance.

Summary: Overall, we find that accounting for the system's bandwidth results in the degrade of the performance of all coding schemes. In particular, different from the mean depth error results in the power-limited setting described in Section S. 4, the Hamiltonian coding schemes outperform sinusoid-based codes at all SNR levels. Although, the choice of best performing coding scheme highly depends on the hardware constraints, practical coding schemes based on Hamiltonian codes seem to consistently be the better choice over sinusoid-based codes.



Supplementary Figure 6: **Mean depth error results for band-limited coding schemes with $f_{max} = 10f_0$.**

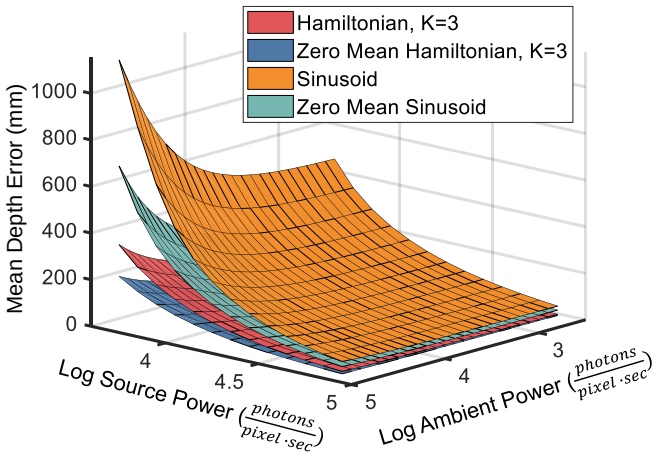


Supplementary Figure 7: **Mean depth error results for band-limited coding schemes with $f_{max} = 5f_0$.**

S. 7. Zero-mean Complementary Demodulation Functions

In this section we discuss one additional important constraint that should be considered when designing practical CW-ToF coding schemes.

Most commercial CW-ToF cameras, if not all, implement a two/four-bucket pixel architecture to efficiently use the total modulation function power [5, 11]. The efficiency of that architecture derives from the fact that any signal arriving at the sensor when the demodulation function is 0 is being discarded. A two-bucket pixel architecture overcomes this by taking two intensity measurement simultaneously where the demodulation functions are negated versions of each other (i.e. complementary functions). Commonly used coding schemes such as 4-bucket sinusoid leverage this pixel architecture to emulate a *zero-mean demodulation function* by taking the difference of these simultaneous measurements. Consequently, a CW-ToF camera based on 4-bucket sinusoid coding only needs to take 2 zero-mean intensity measurements to solve for depths. For the purposes of this discussion we assume sinusoid coding is implemented with low-pass filtered square functions, which is in fact what is done in practice. For more details on the implementation details of sinusoid coding on two-bucket pixel architectures please refer to [5, 11, 7]. Figure 8 shows the mean depth error results for a 4-bucket sinusoid coding implementation that uses a zero-mean demodulation vs. a regular sinusoid coding implementation. Effectively, the regular sinusoid implementation uses half of the total power available decreasing its overall performance.



Supplementary Figure 8: **Performance of zero-mean demodulation vs. regular demodulation functions.** The energy efficient zero-mean demodulation coding schemes outperform their regular implementations. A regular implementation of a coding scheme discards all the photons arriving at the sensor when the demodulation function is 0. For both sinusoid and Hamiltonian $K = 3$ coding schemes this results in only using half of the total power and consequently their performance is significantly degraded. Interestingly, the regular implementation of Hamiltonian $K = 3$ still outperforms the zero-mean sinusoid coding scheme despite using half of the total power.

Zero-mean demodulation constraint: The demodulation functions in practical coding schemes should be able to leverage a two-bucket pixel architecture. This introduces two additional constraints: the number of measurements K should be even, and for every demodulation function there should be a second demodulation function that is the negation/complementary of it.

Zero-mean demodulation Hamiltonian coding schemes: At first glance, none of the practical Hamiltonian coding schemes proposed in this paper adhere to these constraints. However, the binary structure of the demodulation functions makes it simple to extend these coding schemes such that they adhere to the above constraint. For all of the proposed binary coding schemes we can simply double the number of measurements K, where the demodulation functions in the additional K measurements will be the negated versions of the first K demodulation functions. The proposed zero-mean Hamiltonian codes will require $2K$ measurements as opposed to K in their regular counterparts. Figure 8 shows the performance improvement of a zero-mean Hamiltonian $K = 3$ over its regular implementation. It is important to clarify that in this case, zero-mean Hamiltonian $K = 3$ actually had $K = 6$ intensity measurements, however, for the sake of consistent notation we kept its original label.

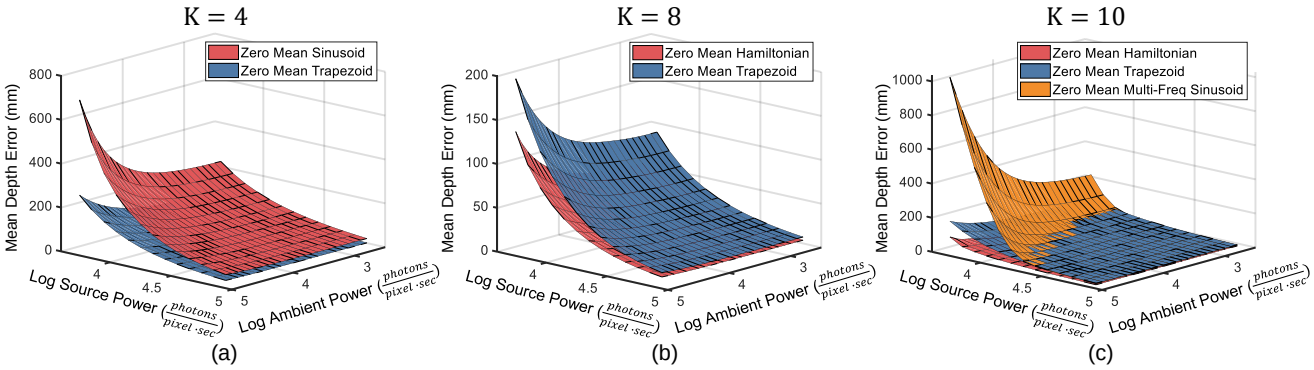
K-bucket trapezoidal coding family: One drawback of the zero-mean demodulation Hamiltonian coding family is that the coding scheme with the least number of measurements need $K = 6$, which translates to 3 measurements with a two-bucket

architecture. On the other hand, zero-mean sinusoid can be implemented with only $K = 4$ measurements. To this end, we propose a new CW-ToF coding family inspired on the structure of Hamiltonian $K = 3$ which we denote as the *trapezoidal coding scheme* family. A trapezoidal coding scheme will have an even number of measurements K, and the modulation and demodulation functions will have the following simple structure:

1. **Modulation Functions:** All K modulation functions are equal. They are square functions with a peak power of KP_{ave} and a duty cycle of $\frac{1}{K}$.
2. **Demodulation Functions:** All demodulation functions are square functions with a $\frac{1}{2}$ duty cycle. Each demodulation function, $D_i(t)$, will have a phase shift equal to $\frac{2\pi}{K}$ radians.

Performance of practical zero-mean demodulation coding schemes: Figure 9 shows the performance of different zero-mean demodulation coding schemes for various K. All simulation parameters are the same for all coding schemes, including total integration time. We omit the comparison of zero-mean Hamiltonian and trapezoidal for $K = 6$ since they correspond to the exact same coding scheme. Similar to zero-mean Hamiltonian, trapezoidal outperforms sinusoid-based schemes at low and medium SNR. Furthermore, zero-mean Hamiltonian codes outperform Trapezoid coding schemes. However, the comparisons in Figure 9 are under ideal hardware settings. Hence, it is important to note that trapezoidal coding schemes have lower peak power and bandwidth requirements than Hamiltonian. Therefore, a performance analysis at different maximum power and frequency levels is required to be able to determine which coding scheme has better performance.

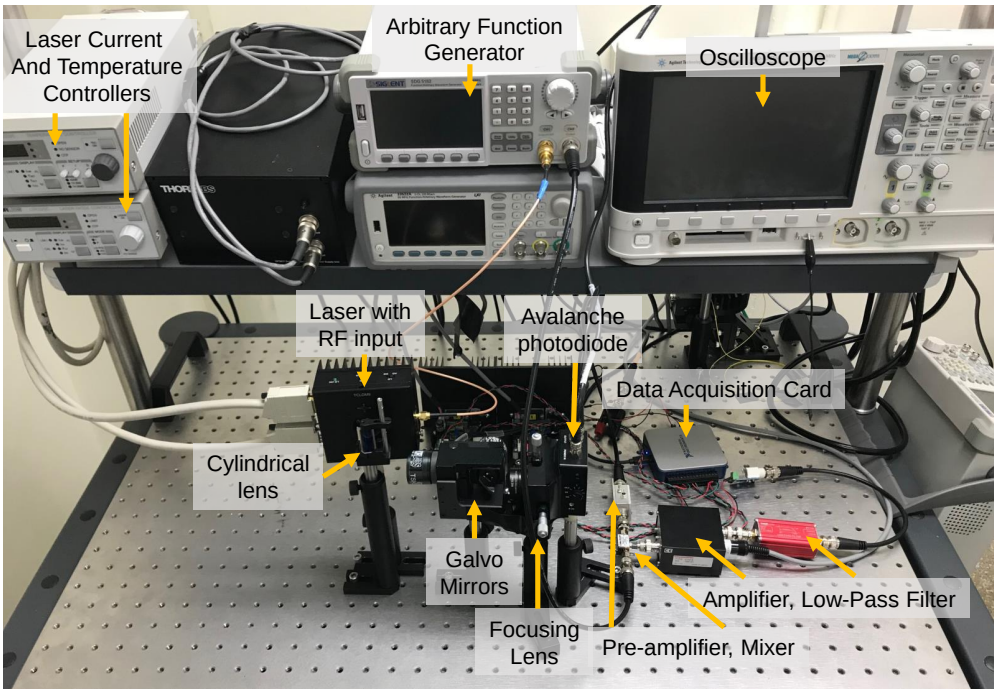
Summary: The zero-mean demodulation constraint described in this section presents an interesting direction of future work on practical coding function design. High-performance coding schemes that adhere to the binary constraint, such as the ones introduced in this paper, can be easily adapted to leverage a two-bucket pixel architecture. Additionally, we propose a Trapezoid coding family that adheres to the zero-mean demodulation constraint and significantly outperforms sinusoid-based coding schemes. Furthermore, they have less strict maximum power and bandwidth requirements than Hamiltonian codes. An analysis under various hardware constraints, similar to the analysis done in this paper, could help determine under which conditions the Trapezoid coding family can outperform Hamiltonian codes.



Supplementary Figure 9: **Performance of zero-mean demodulation coding schemes.** Figure (a), (b), and (c) show the performance of different zero-mean demodulation coding schemes under the same simulation parameters and in an ideal hardware setting (infinite maximum power and infinite bandwidth). The 4-bucket trapezoid outperforms 4-bucket sinusoid (a). In (b) and (c) Hamiltonian codes outperform all coding schemes. However, trapezoid codes have lower maximum power and bandwidth requirements than Hamiltonian.

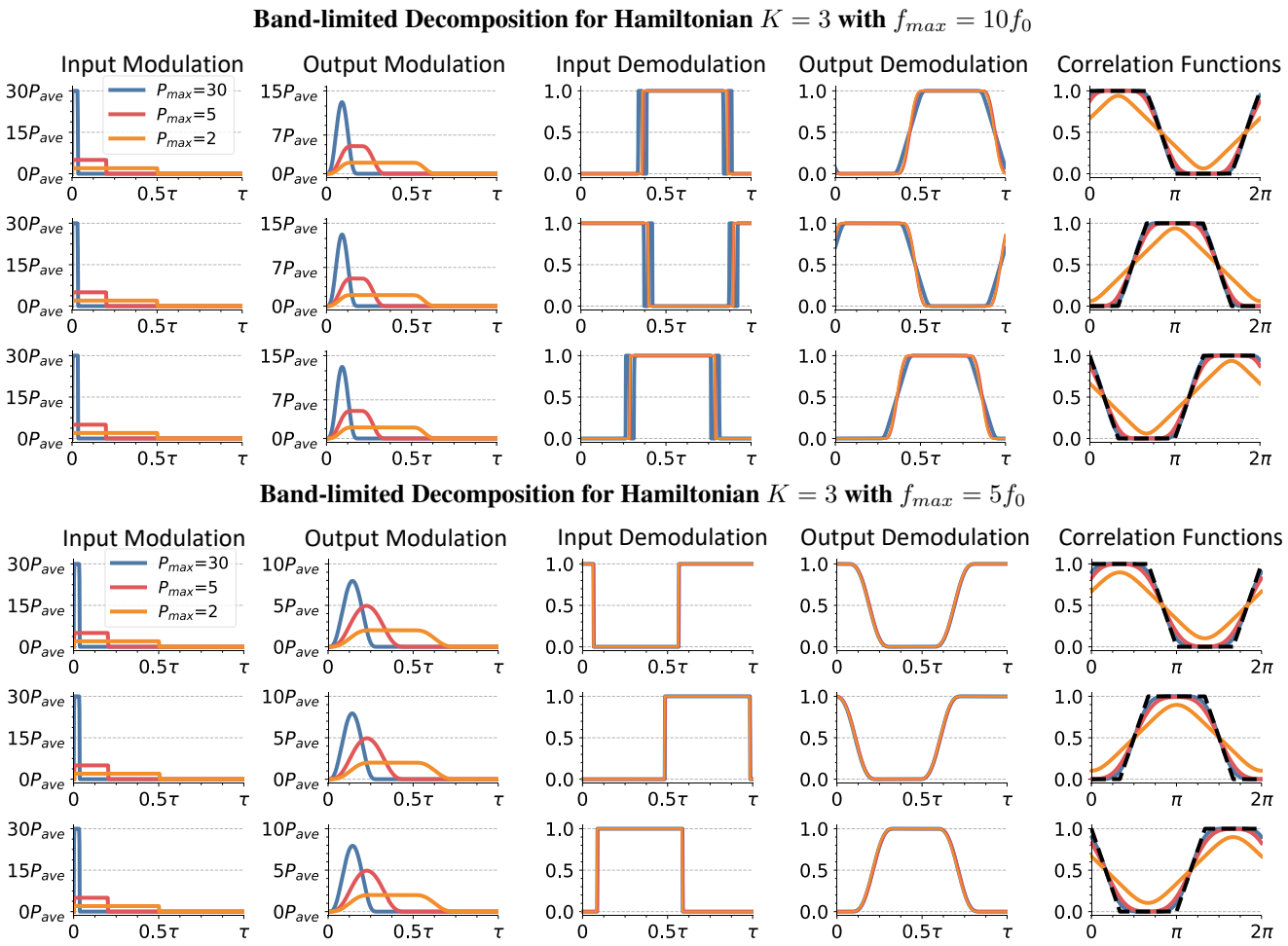
S. 8. Hardware Prototype

Figure 10 shows the CW-ToF single-pixel camera prototype developed to validate the proposed coding schemes. The scene is flash illuminated by the laser diode. A cylindrical lens is used to improve the uniformity of the illumination on the scene. The galvo mirrors scan every point in the scene and the incoming signal is focused onto the APD. The detected signal is multiplied with the demodulation function with a frequency mixer. Finally, the mixer output is amplified, low-pass filtered, and digitized by a data acquisition card, resulting in an intensity measurement.

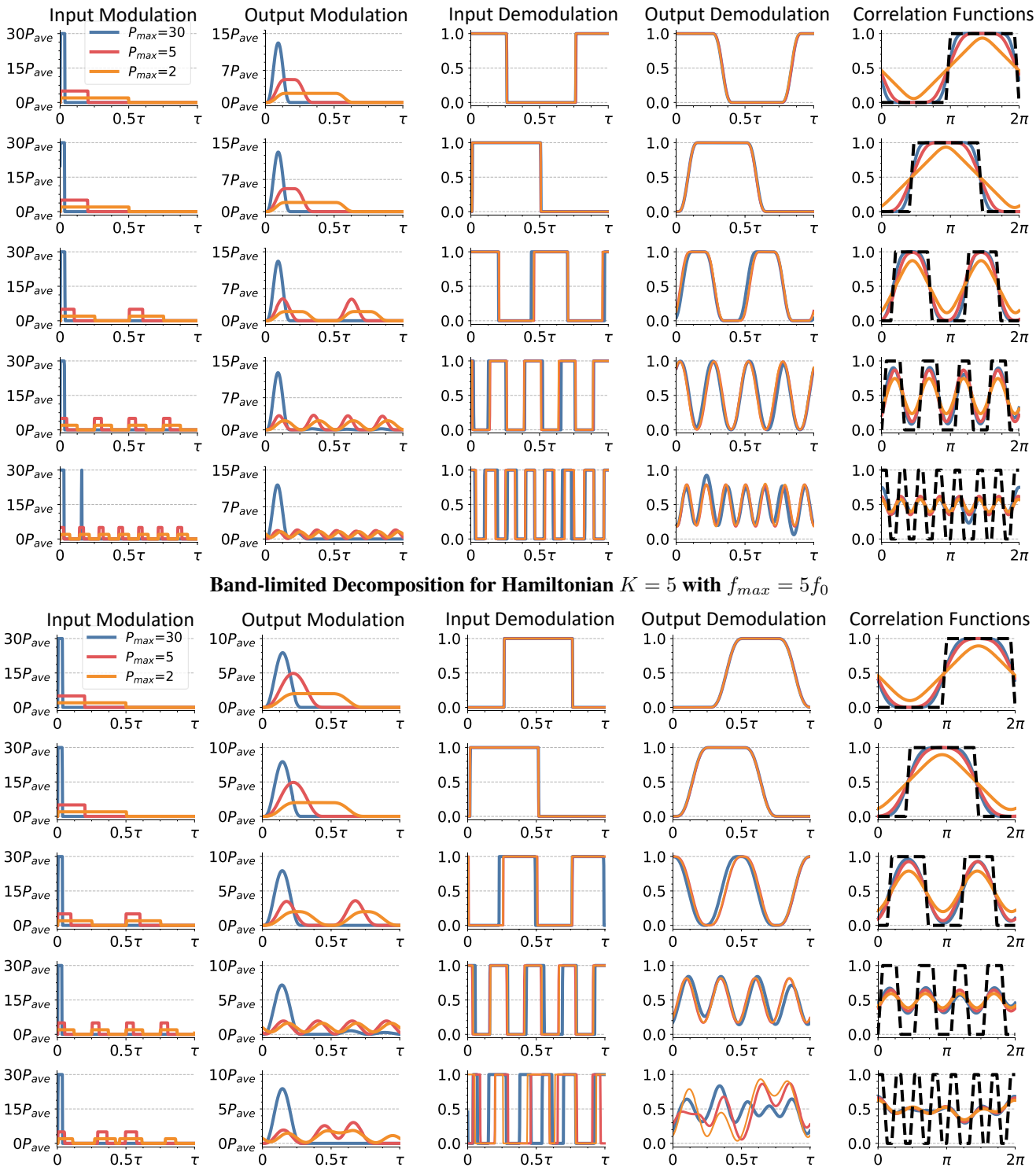


Supplementary Figure 10: **CW-ToF Hardware Prototype.** CW-ToF hardware prototype used for the experiments presented in the paper. The scene is flash illuminated with the laser diode and the galvo mirror system scans the scene onto the APD. The intensity of the laser is modulated with the arbitrary function generator. The detected signal is pre-amplified and then correlated with the demodulation function through a frequency mixer. Finally, the output of the mixer is amplified, low-pass filtered, and digitized, resulting in a intensity measurement.

S. 9. Appendix: Additional Band-limited Correlation Decompositions



Supplementary Figure 11: **Band-limited decomposition for Hamiltonian $K = 3$.**



Supplementary Figure 12: **Band-limited decomposition for Hamiltonian $K = 5$.**

References

[1] A. Adam, C. Dann, O. Yair, S. Mazor, and S. Nowozin. Bayesian time-of-flight for realtime shape, illumination and albedo. *IEEE transactions on pattern analysis and machine intelligence*, 39(5):851–864, 2017. 1

[2] F. Frishman. On the arithmetic means and variances of products and ratios of random variables. Technical report, ARMY RESEARCH OFFICE RESEARCH TRIANGLE PARK NC, 1971. 2

[3] M. Gupta, A. Velten, S. K. Nayar, and E. Breitbach. What are optimal coding functions for time-of-flight imaging? *ACM Transactions on Graphics (TOG)*, 37(2):13, 2018. 1, 2, 3

[4] S. W. Hasinoff, F. Durand, and W. T. Freeman. Noise-optimal capture for high dynamic range photography. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 553–560. IEEE, 2010. 2

[5] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE Journal of quantum electronics*, 37(3):390–397, 2001. 11

[6] P. Mirdehghan, W. Chen, and K. N. Kutulakos. Optimal structured light à la carte. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6248–6257, 2018. 1

[7] A. D. Payne, A. A. Dorrington, and M. J. Cree. Illumination waveform optimization for time-of-flight range imaging cameras. In *Videometrics, Range Imaging, and Applications XI*, volume 8085, page 80850D. International Society for Optics and Photonics, 2011. 11

[8] H. Seltman. Approximations for mean and variance of a ratio. http://www.stat.cmu.edu/~hseltman/files/ratio.pdf. 2

[9] R. Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010. 1

[10] G. Van Kempen and L. Van Vliet. Mean and variance of ratio estimators used in fluorescence ratio imaging. *Cytometry: The Journal of the International Society for Analytical Cytology*, 39(4):300–305, 2000. 2

[11] R. Z. Whyte, A. D. Payne, A. A. Dorrington, and M. J. Cree. Multiple range imaging camera operation with minimal performance impact. In *Image Processing: Machine Vision Applications III*, volume 7538, page 75380I. International Society for Optics and Photonics, 2010. 11