

Sensing and Control for Color Consistency of the Xerographic  
Printing Process

A THESIS  
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF MINNESOTA  
BY

Teck Ping Sim

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
Doctor of Philosophy

Perry Y. Li, Adviser

September 2009

© Teck Ping Sim September 2009

## Acknowledgments

I would like to foremost thank my advisor Professor Perry Y. Li for his patient guidance, support and in setting a high bar and rigor for research excellence. I would also like to thank Professor Ahmed Tewfik, Professor Rajesh Rajamani and Professor William Durfee for serving in my doctoral committee and their advice. Throughout my studies, I am also indebted to Dr. Eric Gross from Xerox Corporation for his helpful discussion and suggestions. I would like to thank him for this much needed assistance.

I would also like to express my extreme gratitude to my wife, Zhang Li, for her love, unwavering support and confidence in me. My thank also goes to my lab mates, past and present for many hours of share fun and encouragement. Lastly to my family for their support, love and understanding in the years spent pursuing my doctoral studies.

Finally, I would like to take this opportunity to gratefully acknowledge the financial support from the Departmental Fellowship, University of Minnesota; the National Science Foundation under grant : CMS-0201622 and Xerox UAC grant in supporting my studies at the University of Minnesota.

Teck Ping,  
September, 2009

## Abstract

This dissertation deals with the formulation, design, and implementation of the sensing and control system for maintaining color consistency of a xerographic color printing process. It involves stabilization of the tone reproduction curves (TRC) of all the primary color separations or directly on the color reproduction characteristics (CRC) function. Color prints quality and consistency are sensitive to disturbance induced variations due largely to the complexity of the print process itself and the small margin of perceptual color variations. In this dissertation a process level control is proposed to directly stabilize the printed colors. Implementing such a control system is difficult because while it is desirable to control all the reproducible colors, we have limited sensing and actuation authorities.

To effectively stabilize the printed colors, all reproducible color information (potentially infinitely many colors) needs to be available at each time instance. However, only small number of color patches (3 to 5 color patches) can be printed and measured at every print cycle. Therefore, unlike typical sensing in control system, we need to sample in both time and in the spatial (or color) domains. Additionally unlike typical sensing of temporal-spatial signal, we are constrained by the availability of sensors to measure the spatial (color) samples and we need to sample these spatial samples in as temporally sparse manner as possible. The proposed solution consists of sampling a small number of  $n$  different colors at different time instances, a method known as time-sequential sampling. With time-sequential sampling, we need to decide which colors, in what order and at what time instance to sample? In this dissertation, an algorithm is proposed to design the time-sequential sampling in the *lattice* theoretic framework for  $n$  samples at each sampling instance known as  $TS(n)$  sampling such that we achieved the largest possible inter-sampling time, while effectively resolving the highest frequencies of the temporal-spatial signal in between samples (i.e. avoids aliasing). We verified that the designed  $TS(n)$  sampling approach is *efficient* and performs as well as if we have full color information and has better performance than equivalent classical time-sequential sampling approaches. Specifically, the key contributions in the design of the  $TS(n)$  sampling include; 1) proposed a framework to parameterize the space of all non-aliasing  $TS(n = 1)$  sampling lattices; 2) showed that the space of all possible non-aliasing  $TS(1)$  sampling is compact. Hence an optimization algorithm that guarantees to terminate can be proposed and has been developed; 3) proposed restriction conditions to reduce the computational complexity of the optimization algorithm; 4) showed that when the temporal-spatial signal is not undersampled temporally by more than a factor of 2, if an optimal  $TS(1)$  sampling can be found under this restriction, the optimal lattice sampling will be as good as any unconstrained optimal  $TS(1)$  sampling; and 5) proposed construction of the  $TS(n), n > 1$  sampling design from the designed optimal  $TS(1)$  sampling. The main

benefit of the  $TS(n), n > 1$  sampling design is the ability to tradeoff sensor hardware and inter-sampling time requirements, an important requirement for implementation in practical system where the inter-sampling time is typically fixed.

In order to achieve color consistency, we need a control system that can constantly maintain a small error between the printed colors and the desired colors at all times. However, the number of colors to be stabilized in this manner far outnumbered the number of actuators in the print process. The proposed solution finds the control input to the actuators that minimize the perceptual error between the printed colors and the desired colors. This approach is akin to finding the best fit printed colors to the desired colors and therefore, all the reproducible colors are “approximately” similar to the desired colors. With the process under control, any residual deviations can now be accounted for with a feed-forward compensation strategy. This control strategy makes use of the known colors deviation and continuously re-coordinate the input colors to the printer such that the desired output colors are achieved. In addition, control approaches were proposed to coordinate and stabilize the tone reproduction of multiple print systems. Both simulation and hardware-in-loop experimental studies were conducted and presented in this dissertation to verify the performance and validity of the proposed approaches.

# Contents

<b>Contents</b>	<b>iv</b>
<b>List of Tables</b>	<b>ix</b>
<b>List of Figures</b>	<b>x</b>
<b>List of Notations</b>	<b>xviii</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Sensing Issues . . . . .	2
1.2 Control Issues . . . . .	4
1.3 Organization . . . . .	6
1.4 Contributions . . . . .	8
<b>Chapter 2 Background</b>	<b>10</b>
2.1 Xerographic Marking Process . . . . .	10
2.2 Digital Color Printing . . . . .	14
2.2.1 Image processing, $f_{htone} \circ f_{sep}$ . . . . .	15
2.2.2 Xerographic color printing, $f_{comb} \circ f_{mark}$ . . . . .	15
2.2.3 Human color perception, $f_{percept}$ . . . . .	17

2.2.4	Digital color printing modeling . . . . .	19
<b>Chapter 3</b>	<b>Sensing of the Temporal-Spatial Signals</b>	<b>23</b>
3.1	Temporal-Spatial Signals . . . . .	23
3.1.1	Linear static plant with uncertainty . . . . .	25
3.1.2	Disturbance model . . . . .	26
3.2	Time-Sequential Sampling . . . . .	32
3.2.1	Framework of time-sequential sampling . . . . .	32
3.2.2	Spectral content of time-sequentially sampled signal . . . . .	35
3.3	Reconstruction of Time-Sequentially Sampled Signal . . . . .	42
3.3.1	Reconstruction via a Kalman filter . . . . .	42
3.3.2	Analysis of time-sequential sampling and reconstruction via a Kalman filter . . . . .	44
3.4	Case Study . . . . .	47
3.5	Summary . . . . .	48
<b>Chapter 4</b>	<b>Lattice Theoretic Framework for Time-Sequential Sampling</b>	<b>53</b>
4.1	Lattice Theoretic Framework . . . . .	57
4.2	$TS(n)$ Sampling Lattices . . . . .	63
4.3	Parameterizing $TS(1)$ Sampling Lattices . . . . .	66
4.3.1	Compactness of the set of non-aliasing $TS(1)$ sampling lattice . . . . .	75
4.4	$TS(1)$ Sampling and Lattices . . . . .	78
4.5	Summary . . . . .	85
<b>Chapter 5</b>	<b>Design of Time-Sequential Sampling</b>	<b>86</b>
5.1	Optimization of $TS(1)$ Sampling Lattices . . . . .	87

5.2	Non-aliasing $TS(1)$ Sampling Lattice for a Given Grid, $LAT(\mathbf{G})$ . . . . .	88
5.2.1	Check if a given lattice, $LAT(\mathbf{GH}_{LTM})$ is a $TS(1)$ sampling lattice .	89
5.2.1.1	Find integer space ${}^{\alpha}\mathcal{V}_N \subset \mathbb{Z}^N$ . . . . .	89
5.2.1.2	Find integer space ${}^{\beta}\mathcal{V}_N({}^{\ell}\mathbf{z}_s)$ for every ${}^{\ell}\mathbf{z}_s \in {}^{\alpha}\mathcal{V}_N$ . . . . .	91
5.2.2	Check if a $TS(1)$ sampling lattice, $LAT(\mathbf{GH}_{LTM})$ is a non-aliasing sampling lattice . . . . .	95
5.2.3	Algorithm to determine if a given lattice $LAT(\mathbf{GH}_{LTM})$ is a non-aliasing $TS(1)$ sampling lattice . . . . .	97
5.3	Optimization Algorithm . . . . .	97
5.3.1	Illustrated example of optimization process . . . . .	101
5.4	Lattice Restriction Conditions . . . . .	102
5.4.1	Restriction due to the loose packing of the frequency content replicas	102
5.4.2	Restriction due to the equivalence of a lattice to its reflected lattice in the temporal-spatial frequency domain . . . . .	104
5.5	Design of $TS(n)$ Sampling Lattices . . . . .	105
5.6	Application of $TS(n)$ Sampling Design . . . . .	109
5.6.1	$TS(n)$ sampling design for TRC . . . . .	110
5.6.2	$TS(n)$ sampling design for CRC . . . . .	113
5.6.3	Efficiency of the $TS(n)$ sampling design . . . . .	116
5.6.4	Performance of the $TS(n)$ sampling design . . . . .	117
5.7	Summary . . . . .	125
<b>Chapter 6 Color Consistency Control</b>		<b>126</b>
6.1	STAGE-I Color Stabilization Control Strategy . . . . .	128
6.1.1	Linear quadratic with integrator dynamics controller . . . . .	129



6.1.2	Robust static controller . . . . .	133
6.1.3	Determination of the error weighting matrix . . . . .	135
6.1.3.1	Color specific weighting, $\mathbf{W}_{\Delta E^*}$ . . . . .	136
6.1.3.2	Image specific weighting, $\mathbf{W}_\varphi$ . . . . .	138
6.2	STAGE-II Color Feed-forward Control Strategy . . . . .	146
6.3	Simulation and Experimental Results . . . . .	148
6.3.1	TRC stabilization control . . . . .	149
6.3.2	CRC stabilization control . . . . .	159
6.4	Summary . . . . .	165
<b>Chapter 7</b>	<b>Color Coordination Control</b>	<b>166</b>
7.1	Coordinated Stabilization Control Strategy . . . . .	167
7.1.1	Optimal static stabilization and coordination controller . . . . .	168
7.1.2	Robust static stabilization and coordination controller . . . . .	170
7.2	Performance of the Coordination and Stabilization Control . . . . .	172
7.2.1	Experiments . . . . .	174
7.3	Summary . . . . .	176
<b>Chapter 8</b>	<b>Conclusion, Discussion and Future Work</b>	<b>180</b>
8.1	Contributions . . . . .	181
8.2	Future Work . . . . .	184
	<b>Bibliography</b>	<b>186</b>
	<b>Appendix A Colorspace Formulation</b>	<b>192</b>
A.1	Basic Color Science Formulation . . . . .	192

A.2	Colorspace Parameterizations . . . . .	193
<b>Appendix B</b>	<b>Formulation Characterizing Temporal-Spatial Frequency Content</b>	<b>196</b>
B.1	Matrix of Fourier Basis Function . . . . .	196
B.2	Finding the Ellipsoidal Compact Spectral Support, $\Theta(W, \mathbf{U})$ . . . . .	199
<b>Appendix C</b>	<b>Algorithms Used for Finding the Optimal <math>TS(1)</math> Sampling</b>	<b>201</b>
C.1	Example of Optimal $TS(1)$ Sampling Design Proposed by Willis and Bresler	201
C.2	Column-Reduced Hermite Normal Form Algorithm . . . . .	203
C.3	Closest Point Search Algorithm . . . . .	204

# List of Tables

- 5.1 Comparison of different non-aliasing sampling design approaches . . . . . 115
- 5.2 Performance of different sampling approaches . . . . . 118

# List of Figures

2.1	Schematic of a typical monochrome xerographic marking process . . . . .	11
2.2	Processes involved in reproducing a color image using a typical color xerographic printer . . . . .	14
2.3	The separated CMY halftone monochromatic images and resulting overlays to reproduce a digital color print . . . . .	16
2.4	Different system topologies for overlaying the color separations to produce a color print. . . . .	16
2.5	The halftone color printing process based on the Yule-Nielsen Spectral Neugebauer Model (YNSN). In this figure, $\bar{w}_i := w_i + d_{mix,i}$ . . . . .	20
3.1	(a) The time-varying TRC signal, $\omega(t, x_1)$ from a legacy printer at nominal actuator inputs $\mathbf{v}_0$ ; (b) The frequency content of the time-varying TRC with the assumed ellipsoidal spectral content, $\Theta(W, U_1)$ . . . . .	26
3.2	(a) The low-frequency band TRC variations from the nominal TRC given by $d_l(t, x)$ at nominal actuator inputs $\mathbf{v}_0$ ; (b) Corresponding power spectral density of $d_l(t, x)$ . The low-frequency band is given by the rectangular box region; (c) The high-frequency band TRC variations from the nominal TRC given by $d_h(t, x)$ at nominal actuator inputs $\mathbf{v}_0$ ; (d) Corresponding power spectral density of $d_h(t, x)$ . The high-frequency band can be characterized to lie in an ellipsoidal region for a given threshold value setting for determination of the significant power spectral density (see Appendix B.2). . . . .	28
3.3	$TS(n = 3)$ sampling points (denoted by $\bullet$ ) on a $1 + N = 3$ dimensional temporal-spatial grid. . . . .	33

3.4	Lexicographical and bit-reversed time-sequential sampling sequence with $M_1 = 16$ tones. . . . .	36
3.5	(a) Spatial Nyquist condition is fulfilled making the replications of the original frequency content, $\Omega(f, \mathbf{u})$ (given by the ellipses) in the $p = 0$ hyperplanes that do not overlap the original frequency content (i.e. the baseband); (b) Spatial Nyquist condition is <i>not</i> fulfilled making the replications of the original frequency content, $\Omega(f, \mathbf{u})$ (given by the ellipses) in the $p = 0$ hyperplane that overlap the baseband . . . . .	38
3.6	(a) Temporal Nyquist condition is fulfilled making the replications of the original frequency content, $\Omega(f, \mathbf{u})$ (given by the ellipses) in the $p = \pm 1$ hyperplanes that do not overlap the original frequency content (i.e. the baseband); (b) Temporal Nyquist condition is <i>not</i> fulfilled making some of the replications of the original frequency content, $\Omega(f, \mathbf{u})$ (given by the ellipses) in the $p = \pm 1$ hyperplanes that overlap the baseband . . . . .	39
3.7	(a) The time-varying TRC signal, $\omega(t, x_1)[1 + N = 2]$ from actual experimental data; (b) $TS(1)$ sampling of $\omega(t, x_1)$ . . . . .	40
3.8	(a) Aliasing weights $ Q_{mp_1} $ for the $TS(1)$ sampling shown in Figure 3.7(b); (b) The frequency content $TS(1)$ sampling of time-varying TRC, $\Omega_s(f, u_1)$ . . . . .	40
3.9	Spectral replication pattern $ Q_{mp_1} $ for lexicographical (left) and bit-reversed sequences (right). . . . .	41
3.10	Transformation of a sinusoidal $1 + N = 2$ temporal-spatial ( $u_1 = 4$ [cycles/tone], $f = 0.01$ [Hz]) disturbance input through the three processes in the Kalman filter. From top to bottom are the temporal spectra for a particular tone: 1) original disturbance signal; 2) transformed by $\tilde{\mathbf{F}}\mathbf{P}(k+1)^{-1}\mathbf{B}_c(k)$ ; 3) after filtering by $\mathbf{A}$ ; 4) demodulation by $\tilde{\mathbf{F}}\mathbf{P}(k)\tilde{\mathbf{F}}^{-1}$ . . . . .	45
3.11	Temporal-tonal power-spectral density for various single tone input frequencies ( $u_1$ [cycles/tone], $f$ [Hz]) after the first transformation of the reconstruction filter for the <b>lexicographic</b> (left) and <b>bit-reversed</b> (right) $TS(1)$ sampling. . . . .	46
3.12	Ratio of mean squared reconstruction errors to mean squared deviation from the nominal TRC, $\Phi$ using lexicographical and bit-reversed sequences at various tonal-temporal support frequencies ( $U_1, W$ ). . . . .	48

3.13	Differences of $\Phi$ between bit-reversed and lexicographic sequences at various tonal-temporal support frequencies $(U_1, W)$ . . . . .	49
3.14	Reconstruction error $(\ \bar{\mathbf{e}}_1(k) - \hat{\mathbf{e}}_1(k)\ _2^2)$ as a function of time using lexicographical and bit-reversed sequences at $(U_1, W) = (3.0, 0.31)$ . . . . .	49
3.15	Input TRC (consisting of a typical TRC with various tonal-temporal support frequencies $(U_1, W)$ imposed on it) and its worst case steady state reconstruction, for Kalman filter designed for the lexicographic strategy (dotted line:reconstructed signal,line:input signal). . . . .	50
3.16	Input TRC (consisting of a typical TRC with various tonal-temporal support frequencies $(U_1, W)$ imposed on it) and its worst case steady state reconstruction, for Kalman filter designed for the bit-reversed strategy (dotted line:reconstructed signal, line:input signal). . . . .	51
4.1	An example of two-dimensional point lattice $LAT(\mathbf{V})[1 + N = 2]$ where $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2]$ . The shaded area is the unit cell, $\Upsilon$ with its volume given by $\vartheta[LAT(\mathbf{V})]$ . . . . .	58
4.2	The replicas of the spatial support $\mathcal{S} \subset \mathbb{R}^N$ on $LAT(\nabla)$ . Note that for ${}^\ell h_{22} \geq M_1$ and ${}^\ell h_{33} \geq M_2$ , regardless of what values ${}^\ell h_{32} \in \mathbb{Z}^+ < {}^\ell h_{33}$ takes, the spatial support will never overlap each other. . . . .	71
4.3	The replicas of the spatial support $\mathcal{S} \subset \mathbb{R}^N$ on $LAT(\nabla)$ . Note that for ${}^\ell h_{33} \geq M_2$ , a proper selection of the ${}^\ell h_{32} \in \mathbb{Z}^+ < {}^\ell h_{33}$ is needed, where in (a) no overlapping of the spatial support replicas occurs whereas in (b) a wrong selection of the ${}^\ell h_{32}$ results in overlapping of the replicas of the spatial support. . . . .	74
4.4	The critical packing lattices of the unit ball $\mathbf{B}$ (top figure) density of the lattice packing, $\delta = \text{volume of } \mathbf{B} / \vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})]^{\frac{1}{2}}$ for different $1 + N$ dimension; (bottom figure) volume of critical packing cell, $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})]$ for different $1 + N$ dimension. In particular, for the case of the TRC ( $1 + N = 2$ ), $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})] = 2\sqrt{3}$ and for the case of the CRC ( $1 + N = 4$ ), $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})] = 8$ . . . . .	76

4.5	The power spectral density of $TS(1)$ sampled signal of $1 + N = 2$ temporal-spatial signal given by $\Omega_s(f, u_1)$ in (4.46). In this figure, $M_1$ is selected such that it fulfills the spatial Nyquist condition and $T$ is selected such it go beyond the temporal Nyquist. The black dots indicate the $(m, p_1)$ position where the spectra replicates. Note that at $m = 0$ spatial plane, the replicas occurs only at alternate $M_1$ values due to (4.48). . . . .	80
4.6	Integrating the power spectral density of the replicas at position $(1, 1)$ on the shaded area gives us $V_{11}$ . . . . .	81
5.1	Two forms of overlapped region at each $i$ -th spatial dimension . . . . .	91
5.2	Possible overlapped region configuration of $\mathcal{O}(\ell_{\mathbf{z}_s})$ for $N = 2$ . . . . .	92
5.3	An example of the $(T, M_1)$ optimization search space for the case of the TRC ( $1 + N = 2$ ). The shaded area shows the area where there exists $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$ such that $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ i.e. $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ is non-empty. . . . .	101
5.4	An example of $TS(2)$ sampling with respect to $\mathcal{S} = [0, A_1)$ by extending the $TS(1)$ sampling lattice with respect to $\mathcal{S}_\top$ . In this case note that we have 2 $\mathcal{S}_\top$ uniformly tiled $\mathcal{S}$ i.e. $\mathcal{S} = \mathcal{S}_\top \cup (\mathcal{S} + \mathbf{g}_1)$ . . . . .	106
5.5	Uniform tiling of the spatial domain, $\mathcal{S}$ for $N = 2$ . . . . .	107
5.6	The $M_1^*T^*$ periodic ( $\square$ ) optimal $TS(1)$ sampling, designed with $W = 0.04$ [Hz], $U_1 = 3.0$ [cycles/tonne]; $A_1 = 1$ ; $M_{max,1} = 50$ and $n = 1$ . Also shown are the $M_1^*T^*$ periodic ( $\circ$ ) pseudo bit-reversed $TS(1)$ , and ( $\cdot$ ) lexicographic $TS(1)$ samplings . . . . .	111
5.7	$\Omega_s(u_1, f)$ with (a) optimal $TS(1)$ sampling; (b) pseudo bit-reversed $TS(1)$ sampling (note that each of the frequency replication for the pseudo bit-reversed sequence is weighed by some values of $ Q_{mp} ^2 > 0.01$ ); (c) lexicographic $TS(1)$ sampling . . . . .	112
5.8	$TS(2)$ sampling sequence (the two samples shown by $\cdot, \times$ at each print cycle) (left) and the corresponding frequency content replication for each signal segment (right) . . . . .	113

5.9	(a) The $\text{prod}(\mathbf{M}^*)T^*$ periodic optimal $TS(1)$ sampling sequence, designed with $W = 0.01$ [Hz], $\mathbf{U} = [2.0, 2.0, 2.0]$ [cycles/tone]; $\mathbf{A} = [1, 1, 1]$ ; $\mathbf{M}_{max} = [8, 8, 8]$ and $n = 1$ . (b) The frequency packing projected onto $(f, u_1, u_2)$ frequency plane at $u_3 = 0$ which shows no overlapping of the projected frequency content of $\Theta(W, \mathbf{U})$ . . . . .	114
5.10	Comparison of sensing efficiency . . . . .	116
5.11	The toner image with the reconstructed TRC (REC image) (top); the toner image with the full sampled TRC (FULL image) at a selected print cycle (middle); the $\Delta E_{94}^*$ map comparing the color tone difference between the REC image and the FULL image (bottom) . . . . .	119
5.12	The reconstruction performance ratio, $\Phi_{k \in [l_0, l_0+L]}$ , $l_0 = 500, L = 500$ versus different degree of sensing noise, $\eta$ for different sampling strategies . . . . .	120
5.13	The simulated time-varying TRC, $\bar{\omega}(k) \in \mathbb{R}^{M=39}$ (top figure); the resulting time-varying TRC after printing and sensing, $\bar{\omega}(k) + \mathbf{n}(k) \in \mathbb{R}^{39}$ , where $\mathbf{n}(k) \in \mathbb{R}^{39}$ gives the sensing noise (middle figure); the distribution of the sensing noise (bottom figure) . . . . .	122
5.14	Comparison of TRC sensing performance . . . . .	123
5.15	The distribution of sensing noise of the $L^*$ , $a^*$ and $b^*$ color output of the time-varying CRC . . . . .	124
5.16	Comparing the performance of using the optimal $TS(1)$ sampling, lexicographic $TS(1)$ sampling, 3-fixed point sampling and full sampling in sensing the time-varying CRC . . . . .	124
6.1	Proposed color consistency control strategy . . . . .	127
6.2	LFT representation of linear system model . . . . .	133
6.3	Finding the significant color map, $\varphi^*$ . Here $DWT$ denotes the discrete wavelet transform and $IDWT$ denotes the inverse discrete wavelet transform	139



6.4	(a) The CSF for luminance (Y) and chrominance (Cr,Cb) channels; (b) Relation between the CSF and a 4-level 2D wavelet decomposition. As an illustration, the shaded areas of the CSF function at the vertical and horizontal frequency axis correspond to the $L_R H_C 2$ of the wavelet decomposition. The weighting that must be used for the wavelet coefficients at each sub-band is described by the corresponding horizontal and vertical frequency range in the CSF functions; (c) 2-level 2D wavelet decomposition with embedded human visual system. . . . .	141
6.5	The HVS FIR filter coefficients for the low-pass sub-band range at $\ell = 1$ . .	142
6.6	The original image, $\varphi$ with the corresponding desired color gamut in the printer's $L^*a^*b^*$ colorspace, $\mathcal{C}_\varphi$ (left); the binary significant map at each YCrCb color planes using $\delta_t = 3\sigma, \sigma = \text{std}( \delta_{cc} )$ (middle); the resulting significant color image, $\varphi^*$ with the corresponding desired color gamut in the printer's $L^*a^*b^*$ colorspace, $\mathcal{C}_\varphi^*$ (right) . . . . .	144
6.7	Significant color image at different threshold settings . . . . .	145
6.8	Comparison of the actual TRC (solid line) and prediction (dotted line) using the nominal sensitivity at four different sets of actuator settings for the Cyan (left), Magenta (Middle) and Yellow (right) primary color separations. Data for this identification is supplied by Xerox. . . . .	148
6.9	Experimental xerographic marking setup for TRC stabilization . . . . .	149
6.10	Sensing and control for TRC stabilization . . . . .	150
6.11	$QPI_L$ of TRC error using full and lexicographic time-sequential sampling at different tonal-temporal frequencies support $(W, U_1)$ . . . . .	152
6.12	Response of TRC stabilization control subjected to induced disturbance with; (a) full sampling; (b) bit-reversed $TS(1)$ ; (c) Lexicographic $TS(1)$ . The curve marked with * is the desired TRC . The $QPI_L$ at steady-state, computed with $l_0 = 50, L = 50$ for the full, bit-reversed and lexicographic sampling is given by 0.00981, 0.01054, 0.01257 respectively. . . . .	153

6.13	Response of TRC for Cyan primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal $TS(1)$ sampling. The curve marked with '×' is the desired TRC. The $MSE_L$ at steady-state, computed with $l_0 = 50, L = 50$ are shown at each respective figures. . . . .	155
6.14	Response of TRC for Magenta primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal $TS(1)$ sampling. The curve marked with '×' is the desired TRC. The $MSE_L$ at steady-state, computed with $l_0 = 50, L = 50$ are shown at each respective figures. . . . .	156
6.15	Response of TRC for Yellow primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal $TS(1)$ sampling. The curve marked with '×' is the desired TRC. The $MSE_L$ at steady-state, computed with $l_0 = 50, L = 50$ are shown at each respective figures. . . . .	157
6.16	The reference image (top row images), the printed time-varying image (middle row images) with; (a) no TRC stabilization control; (b) TRC stabilization control using TRC estimate by optimal $TS(1)$ sampling; (c) TRC stabilization control using TRC estimate by 3 fixed point sampling, at print cycle $k = 50$ . The $\Delta E_{94}^*$ map of the $CMY$ image comparing the top row images and middle row images are given at the bottom rows figures . . . . .	158
6.17	Color consistency with (right) and without (left) color consistency control at different print cycles . . . . .	160
6.18	Experimental xerographic marking setup for CRC stabilization . . . . .	161
6.19	Sensing and control for CRC stabilization . . . . .	162
6.20	Comparing CRC stabilization performance using LQ with integral control for STAGE-I and STAGE-II compensation with the case of no stabilization control.	164
6.21	Comparing CRC stabilization performance using robust static control for STAGE-I and STAGE-II compensation with the case of no stabilization control	164

6.22	Comparing CRC stabilization performance using optimal $TS(1)$ sampling and lexicographic $TS(1)$ sampling. The linear quadratic control is used as STAGE-I . . . . .	165
7.1	Stabilization and coordination performance using the robust static control approach of printer 1 and 2 with different uncertainty level given by different settings of $\delta_1 = \delta_2$ . . . . .	174
7.2	Experimental setup for sensing and control for stabilization and coordination of the TRCs . . . . .	175
7.3	Experimental TRC stabilization and coordination performance for (a) cyan, (b) magenta and (c) yellow primary separations with optimal $TS(1)$ sampling of printer 1 and printer 2. The optimal static control is used here with $\mathbf{S} = 10\mathbf{I}$ for imposing coordination. . . . .	177
7.4	Color image printout of printer 1 (left) and printer 2 (right) at different time-step, $k$ for the optimal static control approach <i>with</i> coordination. Note that color prints at $k = 25$ and $k = 50$ are much more closer to each other than with print at $k = 0$ indicating stabilization of the color prints; and color prints at $k = 50$ of both printers are closer than given at $k = 0$ indicating coordination. . . . .	178
7.5	Color image printout of printer 1 (left) and printer 2 (right) at different time-step, $k$ for the optimal static control approach <i>without</i> coordination. Note that, color prints at $k = 25$ and $k = 50$ are much more closer to each other than with print at $k = 0$ indicating stabilization of the color prints; and color prints at $k = 50$ of both printers are mismatch indicating a lack of coordination. . . . .	179
A.1	Parameterizations by projection onto the principal component . . . . .	195
B.1	The convex hull of the region of significant frequency content, $H_{convex}(\mathcal{U})$ (shaded area), the ellipse, $\Theta(W, U_1) \supseteq H_{convex}(\mathcal{U})$ and associated parameters. . . . .	200
C.1	(a) The assumed spectral support, $\Theta$ ; (b) critical packing for spectral support, $\Theta$ given by (a) . . . . .	201

# List of Notations

The  $N$ -tuples variables are denoted by boldface. Let  $\mathbf{x} = [x_1, \dots, x_N]$  and  $\mathbf{y} = [y_1, \dots, y_N]$  be two vector variables and  $\mathbf{a} = [a_1, \dots, a_N]$  and  $\mathbf{b} = [b_1, \dots, b_N]$  be two  $N$ -tuples integer vectors. The following Matlab styled vector operations apply:

$$(i) \mathbf{x} \cdot * \mathbf{y} = [x_1 y_1, \dots, x_N y_N]$$

$$(ii) \mathbf{x} \cdot / \mathbf{y} = [x_1 / y_1, \dots, x_N / y_N]$$

$$(iii) \mathbf{x} \cdot \text{mod } \mathbf{y} = [x_1 \text{ mod } y_1, \dots, x_N \text{ mod } y_N]$$

$$(iv) \mathbf{x} \cdot > \mathbf{y} = [x_1 > y_1, \dots, x_N > y_N]$$

$$(v) \mathbf{x} \mathbf{y}^T = x_1 y_1 + \dots + x_N y_N$$

$$(vi) \sum_{\mathbf{a}=\mathbf{0}}^{\mathbf{M}-\mathbf{1}} = \sum_{a_1=0}^{M_1-1} \dots \sum_{a_N=0}^{M_N-1}$$

$$(vii) \forall i, a_i \leq b_i, [\mathbf{a}, \mathbf{b}] = \{a_1, a_1 + 1, \dots, b_1 - 1, b_1\} \times \dots \times \{a_N, a_N + 1, \dots, b_N - 1, b_N\}$$

$$(viii) \text{prod}(\mathbf{x}) = \prod_{i=1}^N x_i$$

# Chapter 1

## Introduction

The goal of this dissertation is to formulate, design, analyze and implement a sensing and control for color consistency of a xerographic color printing process. Ideally, the xerographic color printing process should have the following characteristics:

1. the ability to consistently reproduce colors despite disturbances on the xerographic printing system and uncertainty of the xerographic printing process.
2. the ability to coordinate with multiple printers.
3. the ability to perform the aforesaid tasks with minimum sensing resources.

Fulfilling these requirements will enable successful development of high quality, reliable and cost effective printers. In recent years, closed loop feedback control has been recognized as critical in achieving these requirements [1, 2]. However due to the lack of suitable sensors and actuators, and the limited onboard computational power, the majority of the closed loop control systems are for subsystem level control or for regulation of operating conditions. Recent development in sensor/actuator technologies and availability of cheap computational power have enabled direct control of image quality on printed image i.e. *process level* print quality control. Fledging work in this area includes color consistency control [3, 4], banding control [5] and color registration [6]. Yet, the xerographic printing process has sufficiently unique features such that new theoretical perspectives are required to fully leverage the potential of process level control. To maintain high fidelity and consistent color reproduction we need to understand the fundamental sensing and control issues for printing and imaging systems and to develop control systems that are both practical and theoretically sound.

A digital xerographic (i.e. laser) color printer can be considered a mapping between the desired image and the printed output image. An important performance criterion for a color printer is that each color in the desired customer image is faithfully reproduced. By ignoring the spatial dimension (such as lines and textures) of print quality and focusing on the issue of consistent color reproduction only, a xerographic color printer can be represented by the color reproduction characteristics (CRC) function,  $CRC : \mathcal{C} \rightarrow \mathcal{C}$ , desired color  $\mapsto$  output-color, where  $\mathcal{C}$  is a 3-dimensional color space<sup>1</sup>. An ideal printer is one in which the  $CRC$  is an identity function. A digital xerographic color printer generates color by printing and overlaying the Cyan, Magenta, Yellow and black (CMYK) separations. The printing of each color separation is characterized by the tone reproduction curve (TRC),  $TRC : \mathcal{T} \rightarrow \mathcal{T}$ , desired-tone  $\mapsto$  output-tone, where the tone **tone** of the separation is the one dimensional representation of the solidness of the primary toner color.

The xerographic printing process is subject to disturbances from many factors including temperature, humidity, material age and variations etc. Several xerographic actuators [7] such as laser power, corotron voltage and development voltages can be used to combat these variations. The goal of the xerographic color control system is to ensure that the CRC is as close to the identity map as possible. A similar, but simpler control problem can be formulated for the printing of each primary color separation. In this case, the control objective is to maintain and stabilize the TRC for each separation [3]. If the manner in which the primary colors are combined is stable and constant, then the output color will also be consistent when the TRC for each separation has been effectively stabilized.

Both the CRC and TRC control formulations pose significant problems for sensing and control. The sensing and control problems are two key issues addressed in this dissertation.

## 1.1 Sensing Issues

In a typical xerographic process, while the print output is potentially high dimensional (i.e. consists of many different tones/colors) sampling is limited to a few (typically  $n = 3$  samples) tone samples at each print cycle. Which tones should be sensed and how often should these tones be sampled? Process sensing for xerography consists in printing small tone patches in the unused areas of the photoreceptor and measuring their densities using the enhanced toner-area coverage (ETAC) sensors<sup>2</sup>. Typically, only 3 to 5 patches are printed every few photoreceptor belt cycles. This amounts to sampling the tone reproduction curve (TRC) at 3 to 5 points once in a while. Increasing the number of test patches increase hardware

---

<sup>1</sup>which can be parameterized using any consistent color coordinates, e.g. CMY, L\*a\*b\*, x-y-z etc.

<sup>2</sup>The ETAC sensor utilizes illumination from a single LED which has been tuned to a specific wavelength and a PIN diode as the detector[8].

needs as well as consumable (toner) and productivity. Traditionally, only three *fixed* digital area coverages at the light, mid and dark tones are used to represent the TRC for tone consistency control [2, 1]. Hence, variations of the TRC at unmeasured tones will not be adequately compensated. In Groff [9, 10] small number of color test patches are sensed to update an inverse print map by using a piecewise linear function approximation. The use of piecewise linear function approximation allows for the parsimonious update. However, the approach is computationally expensive, requires a great deal of user intervention to obtain good results and most importantly gives no indication of when the piecewise linear map should be updated to ensure color reproduction consistency.

In this dissertation, a time-sequential sampling strategy is proposed to increase the utility of the available feedback information. Time-sequential sampling was investigated in the 1980's and 1990's for video and time varying imaging applications (e.g. video and tomography) [11, 12, 13, 14, 15, 16]. For time varying images, time-sequential sampling refers to sampling the image at different spatial locations at different sampling instances. Rastering in television is an illustration. The benefit is that by trading off temporal bandwidth with spatial bandwidths, the temporal bandwidth of the time varying image can be captured faithfully beyond the Nyquist rate as determined by the periodicity of the sampling scheme alone. In another perspective, the sampling rate can be reduced while retaining the same information content. In our context, the spatial dimension is the tone/color dimension for the case of the TRC/CRC respectively. Hence, time-sequential sampling means that at different sampling instances, different tones or different colors are sampled. This maximizes the information from the TRC/CRC samples and allows the time varying CRC or TRC to be captured (and subsequently reconstructed faithfully via a Kalman filter) even when only a small number of samples of the CRC/TRC are available at each instant.

Allebach [11] provided quite a complete analysis of the spectral content of the time-sequentially sampled signal for a *known* time-sequential sampling sequence. The power-spectral density of the time-sequentially sampled signal was shown to be given by the original power-spectral density, aliased by its weighted spatial and temporal frequency translates. Both the frequency translates and the weights depend on the time-sequential sampling sequence. Hence by changing the time-sequential sampling sequence, the degree of aliasing can be changed. This motivates Allebach [12] to study the design of the time-sequential sampling sequence by minimizing the mean-squared reconstruction error for the special case of white ellipsoidal bandlimited stochastic signals with rectangular spatial support. To avoid an infeasible combinatorial search, a suboptimal solution was proposed involving ad-hoc search over a subset of the possible sampling sequences, which was generated using the Chinese Remainder Theorem. Willis and Bresler [13, 14] provided the theoretical background for an optimal

time-sequential sampling design by restricting the time-sequential sampling sequence to lie on a lattice in the temporal-spatial domain. The restriction to lattice patterns was motivated by the ease and simplicity with which lattice theory was applied to the general multidimensional sampling problem [17]. Their proposed optimal time-sequential sampling design framework eliminates the need to pre-specify the spatial and temporal locations for sampling and is extensible to spatial signals of any dimension and spatial support. In [13, 14], they proposed the design of the optimal time-sequential sampling for time-varying tomography i.e.  $1 + N = 2$  temporal-spatial signal. However, the heuristic approach which requires manual “by inspection” check for frequency content packing as demonstrated in Appendix C.1, cannot be easily extendable to higher temporal-spatial dimension. Furthermore in these earlier works, time-sequential sampling is restricted to having a single sample (i.e.  $n = 1$ ) at each sampling instant.

In this dissertation, we are interested in finding the time-sequential sampling approach that makes the best use of the information obtained from the  $n \geq 1$  samples i.e. we aim to find the time-sequential sampling that avoids aliasing of the assumed spectral support and maximizes the time between  $n$  successive samples for sampling the TRC and CRC. This is made possible by exploiting the fine tonal/color resolution capabilities of most sensors. The key idea here is that by over-sampling the spatial domain, we can tradeoff the spatial bandwidths for the temporal bandwidth. Achieving this requires us to determine 1) which order of sample to take? 2) How to choose the inter-sampling time  $T$ ? 3) If each  $i$ -th spatial dimension with spatial range  $A_i$  is discretized into  $M_i$  points, then what should  $M_i \leq \lceil A_i/\varepsilon_i \rceil$  be?, where  $\varepsilon_i$  denotes the maximum sensor resolution of the  $i$ -th spatial dimension and  $\lceil x \rceil$  is the smallest integer  $\geq x$ . In addition, we proposed the use of a spectrophotometer<sup>3</sup> system to *directly* sample the TRC and CRC on paper. Current approach of using an ETAC sensor measures only tone intensity values (not color measurement) and measurement is taken on patches printed on the photoreceptor drum instead of directly on paper.

## 1.2 Control Issues

A typical xerographic process has a small number of actuators (typically  $m = 3$  number of actuators) and the ability to reproduce high number of colors (such as when a colorful image is printed). Thus, how can the high number of colors be best controlled with a relatively small set of actuators? There are some work in the literature in addressing this. Xerox for example, has practiced closed-loop control in copiers and printers for the last 20 years

---

<sup>3</sup>A spectrophotometer is a device that can measure spectral reflectance or transmittance of color as a function of the wavelength. This measurement can then be converted to CIE L\*a\*b\* values for a standard lighting and observer [18].



and a 3 level control architecture was devised and patented in 1995 [19]. Commands are cascaded from level 3 to level 2 and then to level 1. The level 1 control closely regulates the charging, exposure, development and fuser subsystems. Level 2 control is a supervisory level control that determines the set points in level 1 such that performance is maximized. Level 3 control measures the output of the total system and uses the image processing level to compensate slowly varying transfer and fusing disturbances. It can also accommodate specific reproduction behaviors as requested by the customer with feed-forward control. The three level of control provides a system wide architecture for incorporating different controller designs to increase performance and reliability of stabilizing the tone-reproduction curve (TRC). However, there is no basis of implementation such control architecture in the control theoretic perspective since it is the overall color response that matter and not the intermediate subsystems level. Li and Dianat [3] proposed a more straightforward control system which directly uses the xerographic actuators to compensate for the disturbances on the TRC using a robust static controller i.e. equivalent to lumping level 2 and 3 as compared to the previous approach. Fixed TRC sampling is assumed, and the number of tones that are measured is small compared to the dimensionality of the TRC. The proposed robust control ensures that the TRC behaves adequately even at unmeasured tones in the presence of uncertainty. Since only a small numbers of fixed tones are used, the TRC is not fully characterized by actual measurements. Hence the performance at the unmeasured tones is only as good as the model used to describe them. These TRC stabilization approaches can be extended for color stabilization by stabilizing the TRC for each of the CMY color separation. The idea is that if the manner in which the primary colors are combined is stable and constant, then the output color will also be consistent when the TRC for each separation has been effectively stabilized. This however does not account for the disturbances affecting the color composition. Furthermore the number of xerographic actuators available is very limited and the numbers of tones and colors to stabilize are many. This motivates Groff and Khargonekar [10] to propose using the image processing actuators to compensate for the disturbances on the print process. They achieved this by approximating the inverse of the print map that is updated periodically as the print map drifts as a result of various disturbances. To characterize the inverse print map, the multidimensional piecewise linear functions were used. However, as noted in the previous section, inverting the print map this way has many technical difficulties.

In this dissertation, a two-stage control strategy will be used. An inner control loop utilizes the xerographic actuators to regulate the nominal print engine behavior using the optimal control and robust control approaches. The former consider a nominal xerographic plant model and the latter consider a xerographic plant with uncertainty. This is followed by the use of the image processing actuator for additional control to effectively stabilize the

xerographic printing process. The small number of available xerographic actuators ( $m \approx 3$ ) for each color separation and the need to effectively invert the CRC map at each print cycle pose a significant control challenge. In addition to maintaining color consistency for a single printer, we are also interested in maintaining the color consistency across multiple printers. This is particularly critical for tasks requiring printing large quantities of the same image or collection of images with multiple printers. This is an important development for recent technological development in digital color printing called *reconfigurable parallel printing* [20]. In this scenario, maintaining printer-to-printer colors/tones consistency is as important as page-to-page consistency. This is because if the printers are left uncoupled (i.e. maintaining the colors/tones consistency from printers-to-printers is not considered), then prints from one printer to another may appear different although the page-to-page consistency is maintained. By recognizing that it is the coordinated response of both printers that matters, final performance goal can be achieved via the use of coordination control [21].

Identifying opportunities for actuation and sensing for effective xerographic imaging control requires understanding the underlying physics and system architecture trade-offs. In addition, since the final print quality will eventually be evaluated by humans, an understanding of how human perceive images is needed to meaningfully formulate the print quality requirements. This is important in channeling the right control efforts in correcting for print quality defects and in formulating an effective sensing strategy.

### 1.3 Organization

This dissertation is organized as follows:

In Chapter 2 the background materials on the xerographic color printing process and color science essential for the work described in this dissertation will be given. This chapter motivates the need for a color consistency control system and provides fundamental understanding on opportunity for sensing and control.

In Chapter 3 the problem for sensing the high-dimensional time-varying tone reproduction curve (TRC) and color reproduction characteristics (CRC) function under the constraint that only a small number of  $n$  tones/colors samples can be taken at each print cycle is described. The time-varying TRC and CRC can be generally described by the  $1 + N$  dimensional temporal-spatial signals i.e. for the case of the time-varying TRC,  $1 + N = 2$  and for the case of the time-varying CRC,  $1 + N = 4$ . The time-sequential sampling approach is proposed for sensing such temporal-spatial signals. The temporal-spatial signal can be recovered from the time-sequentially sampled signal through a reconstruction process via a periodic Kalman filter.

Chapter 4 and Chapter 5 address the problem of designing effective time-sequential sampling to sample a  $1 + N$  dimensional temporal-spatial signal. Time-sequential sampling that uses  $n \geq 1$ , where  $n$  is small spatial samples at each time instance is denoted here as  $TS(n)$  sampling. The two design objectives for effective  $TS(n)$  sampling are 1) to avoid aliasing and 2) to obtain the largest possible inter-sampling time. These design objectives were achieved by restricting the  $TS(n)$  sampling to lattices. Chapter 4 gives the lattice theoretic framework for designing the  $TS(n)$  sampling. In Chapter 5 a computation procedure is given for designing the optimal  $TS(1)$  sampling that fulfills both 1) and 2). The  $TS(n)$  sampling for arbitrary  $n > 1$  is then constructed from the optimal  $TS(1)$  sampling design. While this  $TS(n)$  sampling construct can be shown to fulfill objective 1), it does not necessarily fulfill objective 2). Hence, the  $TS(n), n > 1$  sampling construct is not optimal. However, this design is of practical importance because the inter-sampling time of the  $TS(n)$  sampling scale linearly with the number of  $n$  spatial samples. Therefore, it is possible to tradeoff sensing time requirements as often face in practical applications (as in sensing the TRC and CRC) with addition of more sensors hardware.

Chapter 6 presents the color consistency control system. The goal of the xerographic color consistency control system is to control the mapping induced by the xerographic process and the image processing process via feed-forward control such that the tone reproduction curve (TRC)/color reproduction characteristics (CRC) function is close to an identity map. A two-stage control strategy is proposed. An inner control loop (**STAGE-I**) utilizes the xerographic actuators to regulate the nominal print behavior. While we expect that **STAGE-I** control significantly improves the tone/color consistency in the presence of disturbances, the small number of xerographic actuators cannot completely eliminate the TRC/CRC deviations from the desired TRC/CRC. By preserving the TRC/CRC range with **STAGE-I** control in addition to maintaining the TRC/CRC to be relatively stable, we can introduce a feed-forward control (**STAGE-II**) that makes use of the image processing actuator to compensate for the residual TRC/CRC variations.

Chapter 7 considers the problem of coordinating and stabilizing two TRCs of two xerographic printing engines. Addressing this problem is critical in tasks requiring printing large quantities of the same image or collection of images with multiple printers, where maintaining printer-to-printer colors/tones consistency is as important as page-to-page consistency. This is because if the printers are left uncoupled, then the prints from one printer to another may appear different although the page-to-page consistency is maintained. In this dissertation an optimal static and robust static controllers are proposed to stabilize and coordinate the potential high dimensional TRC with limited actuation and sensing capabilities. The key idea is to minimize the overall least squares deviation of the two TRCs and the overall

least squares deviation of each of the TRC from the nominal subject to differing degree of disturbances on the printers.

Chapter 8 contains the conclusion and possible directions of future research.

## 1.4 Contributions

1. To *fully* characterize the tone reproduction curve (TRC) and color reproduction characteristics (CRC) function by *actual measurements*, the time-sequential sampling approach initially investigated for video and time-varying imaging applications [11] was employed. The key contributions of this work are given as follows:
  - (i) Here, the TRC and CRC are two and four dimensional temporal-spatial signals respectively. Hence, time-sequential sampling of the temporal-spatial signal refers to the sampling of a small number of  $n$  different tones/colors (i.e.  $n$  number of sensors) at different time instances. The time-sequentially sampled signal is then reconstructed using a periodic Kalman filter to obtain a full characterization of temporal-spatial signal [22]. In [23], the classical time-sequential sampling approach was successfully applied to maintain the tone consistency of the xerographic printing process.
  - (ii) In [24, 25, 26], an optimization approach to design the optimal  $TS(1)$  sampling was proposed to give the largest possible inter-sampling time, while effectively resolving the highest frequencies of the temporal-spatial signal in between samples (i.e. to avoid aliasing) for arbitrary  $1 + N$  temporal-spatial dimension. This work also extends this design approach for any  $n > 1$  samples at each time instance, as oppose to the one sample point with current time-sequential sampling approaches. This design process is made possible by restricting the time-sequential sampling in the lattice framework. Optimal  $TS(1)$  sampling and its extension  $TS(n), n > 1$  sampling is efficient and performs as well as if we had full color/tone information and have better performance than an equivalent classical time-sequential sampling approach.
2. Since, the number of tones/colors to stabilized far outnumbered the number of actuation parameters, a robust static tone reproduction stabilization controller was proposed in [3]. This approach is akin to finding the best fit reproduced (printed) tones to the desired tones subject to print process uncertainty. In this dissertation this approach is extended in three ways which make up the key contributions of our work in this area:

- (i) First, in [27, 28, 23, 29], the curve fitting approach was extended by allowing stabilization of the tone reproduction function i.e. **STAGE-I** that are *fully* characterized by *actual measurements* through the time-sequential sampling approach. In addition, a linear quadratic with integrator dynamic control approach was designed and implemented to allow  $q < m$  tones to be precisely controlled. With the process under control, any residual deviations can now be accounted for with a feed-forward control strategy i.e. **STAGE-II**. This control strategy makes use of the known tone deviations and continuously re-coordinates the input tones to the printer such that the desired output tones are achieved.
- (ii) Secondly, optimal static and robust static control approaches were proposed to coordinate and stabilize the tone reproduction of multiple print systems [30].
- (iii) Finally, the proposed tone stabilization approaches were extended for color reproduction by stabilizing/coordinating each of the primary TRCs [29, 30] and *directly* on the CRC [31][32].

## Chapter 2

# Background

In this chapter the background materials on the xerographic color printing process and color science essential for the work described in this dissertation will be given. The xerographic marking process and the characteristics of this process is discussed in section 2.1. The xerographic marking process reproduces only halftone monochromatic images. By overlaying the CMYK monochromatic images reproduce with the xerographic marking processes, a color print can be realized. The digital color printing process consisting of the composition of the human perceptual model, the xerographic marking process and the related image processing processes is described in section 2.2. In addition, the digital color printing process modeling with the Yule-Nielsen Spectral Neugebauer model is given at the end of this section. Section 2.1 and 2.2 motivates the need for a color consistency control system and provides fundamental understanding on the opportunity for sensing and control.

### 2.1 Xerographic Marking Process

The main component of the xerographic marking process is the chargeable photoconductive insulating surface or photoreceptor. It is usually in the form of a drum or a belt that continuously rotates and cylindrically interacts with several stationary subsystems as shown in Figure 2.1 [7, 33]. Typically, the velocity of the photoreceptor is controlled, as the velocity and the image size determine the print rate. The photoreceptor acts as an image carrier where a toner image is first built up before being transferred to paper, or to an intermediate transfer medium. The density profile of the toner image, which is critical to the print quality, depends on the behaviors of the following process steps:

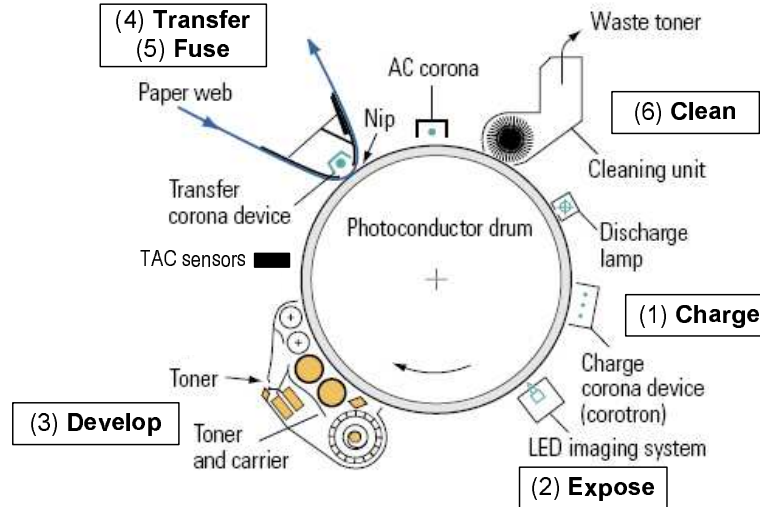


Figure 2.1: Schematic of a typical monochrome xerographic marking process

- (1) **Charge:** A corona discharged from a high voltage corotron wire or a charge roller causes air to breakdown into charged particles. These charged particles uniformly charge the surface of the photoreceptor. The charge level depends on the voltage across and the condition of the corotron wire. To produce more uniform photoreceptor charging, a screen is sometime added between the corotron wire and the photoreceptor. In this case the screen potential will also influence the charge level. The implication of the varying charge level as a consequence of its dependence on the condition of the corotron wire and screen potential, results in varying print quality.
- (2) **Expose:** Light produced by the binary operated raster laser beam scans the photoreceptor line by line to discharge the normally charged photoreceptor. In a Discharged Area Development (DAD) process, light is shone in locations where toners are desired. This creates on the photoreceptor a discharge density profile, known as the latent image, which resembles the desired image. The depth of the discharge is effected by the laser power according to the nonlinear Photo-Induced Discharged Characteristics (PIDC) of the photoreceptor.
- (3) **Develop:** In this stage toner particles are deposited on the photoreceptor. Toner particles are pre-charged by agitating against magnetic carrier beads via the triboelectric process. By virtue of the bias voltage of the development housing and the photoreceptor, the charge toner particles are selectively deposited in the discharged regions (the latent image) on the photoreceptor, transferring it into the desired image. The development process depends heavily on the electric field profile in the development gap, the bias voltages on the development housing, toner charge density (tribo),

as well as toner surface properties such as cohesivity. In most high-end xerographic printers, enhanced toner-area coverage (ETAC) sensors are available to measure the toner mass on the portion of the photoreceptor covered by the toner after the development stage. The ETAC sensing head utilizes illumination by a single LED which is tuned at a specific wavelength and uses a PIN diode as the detector [8].

- (4) **Transfer:** The photoreceptor is now brought in contact with an intermediate medium or a sheet of paper so that the toner particles which were loosely attached electrostatically to the photoreceptor are deposited on the medium. This process is made possible by pre-charging the intermediate medium or the paper to have higher potential relative to the photoreceptor.
- (5) **Fuse:** The image is permanently fixed on the paper by melting the plastic coating of the toner particles on the paper surface with high pressure and temperature.
- (6) **Clean:** The photoreceptor is discharged and cleaned of any excess toner using coronas, lamps, brushes and/or scraper blades. Residual toner will otherwise contaminate the next latent image that the photoreceptor will generate.

In this dissertation, the spatial artifacts (such as lines and textures) of the print quality are ignored. Instead we focused on the consistency of the tone and color reproduction. Because the laser beam in the expose step is only pulsed either on or off on a micro-scale, toner can only be either present or absent. Thus, to reproduce a given partial tone, a halftoning process [34] is needed. In this process, the solidness of the reproduce tone is translated into dot patterns of different densities arranged in a uniform screen pattern. Human eye is spatially low-pass [33] making sparsely (densely) populated dot patterns as being perceived as having light (dark) tones. The tone perception depends on the dot locations and on how each dot absorbs light (“dot gain”). Hence, the xerographic marking process as described here at any time,  $t \in \mathbb{R}^+$  can be represented by the tone reproduction curve (TRC)

$$TRC(t) : \mathcal{T} \rightarrow \mathcal{T}, x_{desired} \mapsto x_{printed} \tag{2.1}$$

where the tone  $x \in \mathcal{T} \subset \mathbb{R}$  of the separation is the solidness of the primary toner color. Typically and as will be used throughout this dissertation,  $\mathcal{T} := [0, 1)$ . In this form, a patch with tone,  $x = 0.1$  for the magenta color separation corresponds to a light violet whereas  $x = 1.0$  corresponds to solid magenta color. Physically, the tone of the primary separations is determined by the pattern and size of the half tone dots printed. Roughly speaking, the denser and bigger the dots are, the more solid the color.



From the description of the xerographic marking process, to maintain the consistency of the TRC map, it is necessary:

1. to generate the same discharge density profile on a print-to-print basis
2. to generate the same electrostatic forces to enable adhesion of the toner particles of the photoreceptor in the development process
3. to generate the same electrostatic forces to transfer the toner particles from the photoreceptor to paper or an intermediate medium
4. to consistently clean all the residual toner from the photoreceptor

Unfortunately, it is difficult to achieve these objectives because even the most well designed xerographic marking process is subject to uncontrolled changes in the material states of the consumable (i.e. toner, photoreceptor, media, etc.) and environmental factors. Environmental factors refer to the ambient operating temperature, relative humidity, and altitude. The time scales of these state changes vary considerably. For example, the multi-layer organic photoreceptor which degrade with time results in much slower changes in the printout compare to the effect of changes in the environmental factors. In summary, the xerographic process is highly interactive with multiple inputs and multiple outputs, and subject to slowly varying and often, not measurable changes in the operating conditions. Furthermore, it is usually difficult if not impossible to characterize the effects of these changes on the print quality. Thus, regulating the subsystems of the xerographic engine alone will not be sufficient to maintain the tone consistency.

Fortunately several parameters in the xerographic process can be adjusted and can therefore be used for xerographic control to maintain the consistency of the TRC map. They include: the voltage across the corotron wire and the screen voltage of the charging system which affects the charged photoreceptor voltage; laser power which affects the discharged photoreceptor voltage; and the bias voltage in the development system which impacts the toner deposition on the photoreceptor. Typically, about  $m = 3$  to 5 xerographic actuators are available.

For  $i$ -th primary separation, let  $\varphi_i(\mathbf{v}_i, \mathbf{d}_i)[x]$  denote the toner area coverage when the marking process prints a halftone image of tone  $x \in [0, 1)$  using  $\mathbf{v}_i \in \mathbb{R}^{m_i}$  as the settings for its  $m_i$  actuators; and  $\mathbf{d}_i$  gives the unknown disturbance on the xerographic printing process. Although the TRC is infinite dimensional, it can be adequately approximated by its response at a finite number of tones  $x_1, \dots, x_M$ , where  $M \gg 1$  can be large. This is possible since

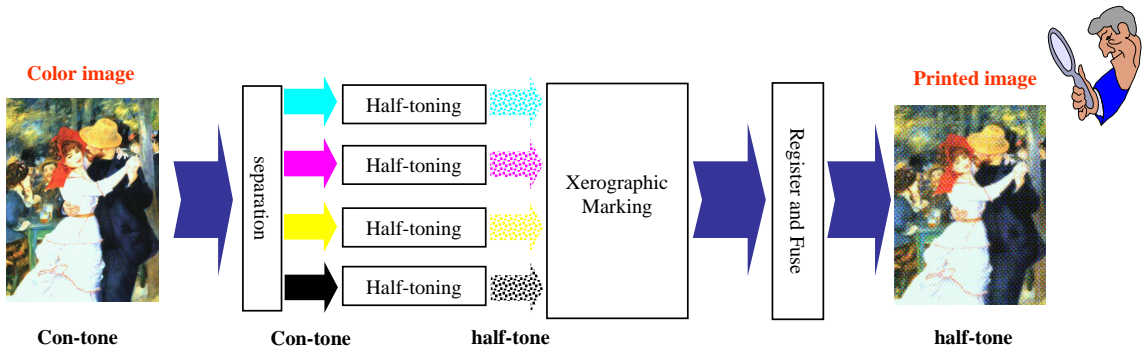


Figure 2.2: Processes involved in reproducing a color image using a typical color xerographic printer

the human visual system cannot differentiate tones that are close together (with  $\Delta E_{94}^* \leq 1$ , see section 2.2.3). Thus, the xerographic control process can be represented by:

$$TRC(kT) := \varphi_i(\mathbf{v}_i, \mathbf{d}_i) = \begin{pmatrix} \varphi_i(\mathbf{v}_i, \mathbf{d}_i)[x_1] \\ \vdots \\ \varphi_i(\mathbf{v}_i, \mathbf{d}_i)[x_M] \end{pmatrix} \in \mathbb{R}^M \quad (2.2)$$

The time index  $k$  typically corresponds to the number of photoreceptor belt revolutions because measurements using the TAC sensors are available once every belt cycle, with inter-sampling time,  $T$ . A nominal TRC is generally designed so that it can be achieved when no disturbance is present. Let the nominal TRC be given by  $\varphi_i((\mathbf{v}_0)_i, \mathbf{0})$  where  $(\mathbf{v}_0)_i \in \mathbb{R}^{m_i}$  is the nominal control of the  $i$ -th primary separation. Although the nominal plant behavior can be obtained experimentally and through physical modelling in the product design process, the precise effect and indeed the sources of disturbances cannot be easily determined. The nonlinear model in (2.2) is therefore not well known.

## 2.2 Digital Color Printing

The color print quality of the digital color printing as illustrated in Figure 2.2 can be represented by the color reproduction characteristics (CRC) function.

$$CRC(t) : \mathcal{C} \rightarrow \mathcal{C}, \mathbf{x}_{desired} \mapsto \mathbf{x}_{printed} \quad (2.3)$$

where color  $\mathbf{x} \in \mathcal{C} \subset \mathbb{R}^3$  is a vector that uniquely quantifies a given color for standard viewing and lighting condition. The CRC is potentially infinite dimensional because ideally any specified colors can be reproduced.

Producing a digital color print involves three main processes : *image processing* (software), *xerographic marking* (hardware) and *human perception* (psychophysics) (see Figure 2.2).

Hence,

$$CRC(t) := \underbrace{f_{percept}}_{\text{human}} \circ \underbrace{f_{comb} \circ f_{mark}}_{\text{printer}} \circ \underbrace{f_{htone} \circ f_{sep}}_{\text{image processing}} \quad (2.4)$$

where  $f_{sep}$  separates the image into primary color planes,  $f_{htone}$  performs half-toning on each color separations,  $f_{mark}$  prints the half-tones separations with toner,  $f_{comb}$  forms a composite image from the printer separations, and  $f_{percept}$  models the human color perception.  $f_{sep}$  and  $f_{htone}$  are carried out in the image processing stage;  $f_{mark}$  and  $f_{comb}$  are the physical xerographic process and  $f_{percept}$  resides in the human psychology and physical stimulus.

The added complexity and uncertainty of the physical printing process and related image processing processes, and the sensitivity of the human visual system to color variations [18], makes it difficult to produce high fidelity and consistent color prints. The problem is however not insurmountable because both the *image processing*,  $f_{sep}$  process and small number of actuators in *xerographic marking*,  $f_{mark}$  process can be used to compensate for the print quality variations.

### 2.2.1 Image processing, $f_{htone} \circ f_{sep}$

The *image processing* process involves separating the desired color image into typically four monochromatic images i.e. the Cyan, Magenta, Yellow and black (CMYK) monochromatic images of varying tonal densities depending on the overall image to be produced. The separation process is typically achieved by using the inverse map of the actual printing process. To improve the print quality, in current printing system this inverse map is updated through the device characterization methods [34]. The separated monochromatic images are reproduced separately with the halftoning process [34]. Note that in principle the C, M and Y colorants suffice to produce all perceivable colors. The black, K colorant is used to replace grey-scale colors (i.e. when C=M=Y) because it is considerably less expensive [34]. Therefore for our purpose it suffices to consider only the CMY separations.

Throughout this dissertation we assume  $f_{sep}$  as an identity map unless explicitly stated otherwise. This means that  $CRC(t)$  in (2.4) is a map with input CMY colorspace.

### 2.2.2 Xerographic color printing, $f_{comb} \circ f_{mark}$

In the xerographic color printing process, the CMY halftone monochromatic images shown in Figure 2.3 can now be produced separately and overlaid on one another. Thus, in this case the *xerographic marking* process described in the previous section has to be repeated 3 times for a single color print. For example, in a multi-pass topology (see Figure 2.4(c)),

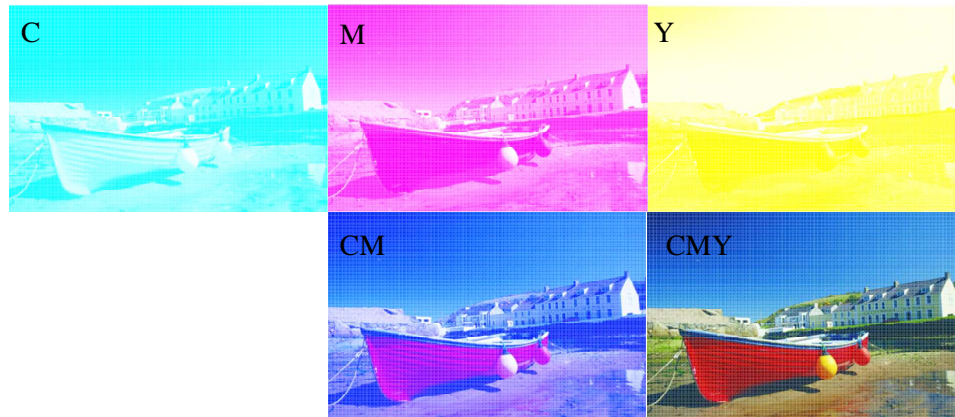


Figure 2.3: The separated CMY halftone monochromatic images and resulting overlays to reproduce a digital color print

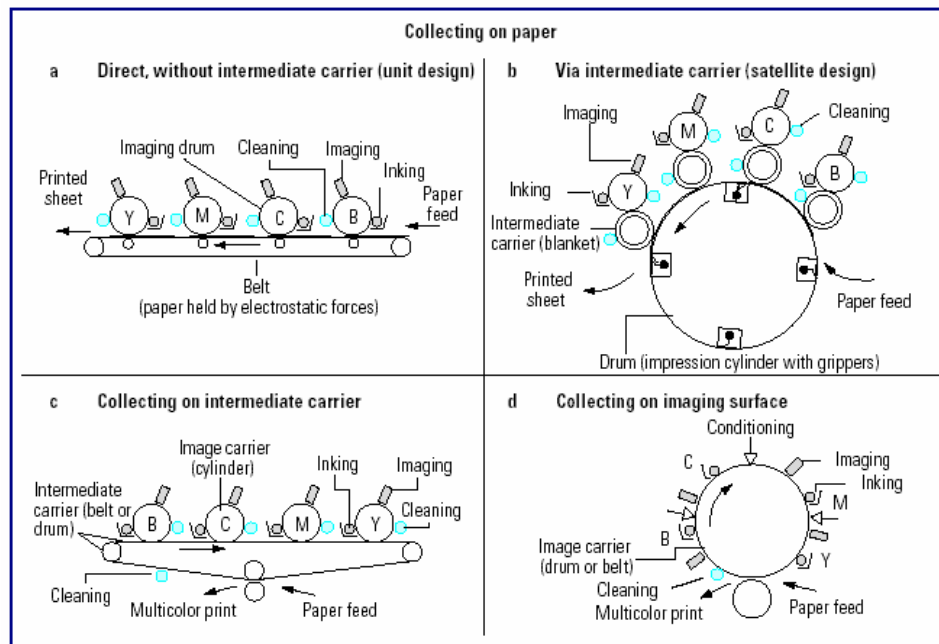


Figure 2.4: Different system topologies for overlaying the color separations to produce a color print.

for each color separations, the photoreceptor has to be charged, exposed, and developed, and the developed toner image has to be transferred to an intermediate medium. Once all the color separations have been laid down on the intermediate medium, the composite image is then transferred to the paper, and fused all at once. Other topologies as shown in Figure 2.4 produce a different level of productivity, require a different amount of hardware duplication, and have different sensitivities to disturbances.

### 2.2.3 Human color perception, $f_{percept}$

In developing the proposed color consistency control strategy we need to exploit features of human visual perception models for performance analysis, control synthesis and effective sensing strategy. This allows the control and the sensing systems to focus their resources and efforts on the aspect of image quality that is most relevant to the human observer.

The perception of color is the result of interaction between a physical stimulus, the eye's receptors and the brain. The eye's receptors sense the stimulus and the sensed signal is sent to the brain for interpretation. The physical stimulus for color is light. The human eye has three types of cone cells called L, M and S cones, which are responsible for perceiving color. L, M, and S correspond to the Long, Medium and Short wavelength sensitive cones respectively. Under fixed set of viewing conditions, the response of these cones can be modeled by a linear system defined by the spectral sensitivities of the cones

$$c_i = \int_{\lambda_{min}}^{\lambda_{max}} s_i(\lambda)L(\lambda)\mathbf{r}(\lambda) d\lambda \quad i = 1, 2, 3 \quad (2.5)$$

where  $s_i(\lambda)$  denotes the sensitivity of the  $i$ -th type cones,  $L(\lambda)\mathbf{r}(\lambda)$  denotes the spectral distribution of light incident on the retina, and  $\lambda_{min}, \lambda_{max}$  denote the interval of wavelengths outside of which all the sensitivities are zero. Mathematically, (2.5) correspond to inner product operations in the Hilbert space of square integrable functions,  $\mathcal{L}^2([\lambda_{min}, \lambda_{max}])$ . Hence, the cone respond mechanism corresponds to a projection of the spectrum onto the space spanned by three sensitivity functions  $\{s_i(\lambda)\}_{i=1}^3$ . This space is called the *human visual subspace*.

This forms the basis of colorimetry, which represents all perceivable colors by three independent coordinates. The three values vector defined by this coordinate system is called the tristimulus values. For computation, the spectral quantities in (2.5) may be replaced by their sampled counterparts to obtain summations as numerical approximations to the integrals. The International Commission on Illumination,(CIE) provides standard sampled values of illuminant,  $L(\lambda)$  and observer spectral sensitivity functions,  $s_i(\lambda)$  [18] that can be

used to compute the CIE XYZ tristimulus values for a measured reflectance,  $\mathbf{r}(\lambda)$  of a given color.

However the XYZ colorspace is *not* perceptually uniform in that equal Euclidean distance between points in the XYZ colorspace at different regions of the color space does not give equal perceptual differences. Because perceptual uniformity is an extremely desirable feature for defining tolerances in color reproduction systems, a perceptually uniform colorspace is desirable. A “nearly” uniform colorspace was defined by CIE in 1976 [18] known as the CIE  $L^*a^*b^*$  colorspace. The  $L^*a^*b^*$  colorspace is obtained from a nonlinear transformation of the XYZ colorspace. Since the  $L^*a^*b^*$  and XYZ colorspace are based upon some fixed standard, it does not depend on the device that reproduced a given color specifies by the  $L^*a^*b^*$  or XYZ values. Hence the  $L^*a^*b^*$  and XYZ colorspace are what is known as the device independent colorspace. In contrast, colorspace such as the CMY (Cyan, Magenta and Yellow) for printing and RGB (Red, Green and Blue) for display are known as the device dependent colorspace. The colors they represent depend on the specific primary color pigments or light of the device that reproduce them and on the settings and configurations of that device. Therefore, it is not meaningful to specify colors in the device dependent colorspace since they are not indicative of the actual colors viewed by a human observer. Worst still, it is impossible to compare the performance of different print system using this representation. This makes the color independent colorspace truly essential for this work.

To effectively quantify the color difference of any two colors, the  $\Delta E_{94}^*$  color difference formulation is used. The  $\Delta E_{94}^*$  is a generalized color-difference equation using CIE  $L^*a^*b^*$  as an approximately uniform colorspace, developed based a number of color psychophysics experiments by the CIE in 1994. The  $\Delta E_{94}^*$  value gives a better color difference quantification compared to just taking the 2-norm of difference called the  $\Delta E_{ab}^*$  of two colors in the  $L^*a^*b^*$  colorspace. Let  $L_1^*a_1^*b_1^*$  and  $L_2^*a_2^*b_2^*$  denote the values of two colors in the  $L^*a^*b^*$  colorspace. Then  $\Delta E_{ab}^*$  and  $\Delta E_{94}^*$  are given as follows [18, 34]:

$$\Delta E_{ab}^* = \left[ (L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2 \right]^{\frac{1}{2}} \quad (2.6)$$

$$\Delta E_{94}^* = \left[ \left( \frac{(L_2^* - L_1^*)}{k_L S_L} \right)^2 + \left( \frac{\Delta C_{ab}^*}{k_C S_C} \right)^2 + \left( \frac{\Delta H_{ab}^*}{k_H S_H} \right)^2 \right]^{\frac{1}{2}} \quad (2.7)$$

where

$$\Delta C_{ab}^* = (a_2^{*2} + b_2^{*2})^{\frac{1}{2}} - (a_1^{*2} + b_1^{*2})^{\frac{1}{2}} \quad (2.8)$$

$$\Delta H_{ab}^* = \left[ (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2 - \Delta C_{ab}^{*2} \right]^{\frac{1}{2}} \quad (2.9)$$

$$S_L = 1 \quad (2.10)$$

$$S_C = 1 + 0.045C_{ab}^* \quad (2.11)$$

$$S_H = 1 + 0.015C_{ab}^* \quad (2.12)$$

$$C_{ab}^* = (a_1^{*2} + b_1^{*2})^{\frac{1}{2}} \quad (2.13)$$

$k_L, k_C, k_H$  are positive real-valued scaling parameters. As noted in [18],  $k_L = k_C = k_H = 1$  at reference conditions. The reference conditions assume that a pair of nontextured specimens in edge contact are placed on a background with  $L^* = 50$ , illuminated by D65 with an illuminance of 1000 lux, and viewed at a distance so that the pair subtends a visual angle of at least  $4^\circ$ . Although not explicitly stated, a perceptibility judgment is assumed.

It turns out that the human visual system cannot differentiate any two colors with a  $\Delta E_{94}^*$  of less than one. This tolerance of the human visual system known in the literature as the *just noticeable difference* permits a small margin of error in reproducing a color image. Therefore, we can justify our approach in stabilizing a finite collection of  $L^*a^*b^*$  values corresponding to perceptually significant colors in order to achieve print quality consistency. The perceptual model of colors embedded in the  $L^*a^*b^*$  colorspace meaningfully formulates the print quality and sensing requirements.

The basic color science formulations mentioned here and used through this dissertation are given in Appendix A.1.

#### 2.2.4 Digital color printing modeling

The digital color printing process can be modeled by the Yule-Nielsen Spectral Neugebauer (YNSN) model as shown in Figure 2.5. In the YNSN model, the macro-reflectance of a colored halftone is viewed as  $w_i$  weighted average of the micro-reflectance of the individual primary reflectance known as the Neugebauer primary reflectance,  $\mathbf{r}_{P,i}(\lambda)$ . The Neugebauer primary reflectance are made up of the overlay of the CMY halftone monochromatic images. This result in  $2^3 = 8$  different primaries through subtractive overlap of one, two, three or no colorants. To account for the penetration and scattering of light on paper, the Yule-Nielsen factor,  $n_y$  is incorporated in the model. The model gives the resulting predicted reflectance  $\mathbf{r}(\lambda)$  of the printed halftone color. The YNSN equation is given as

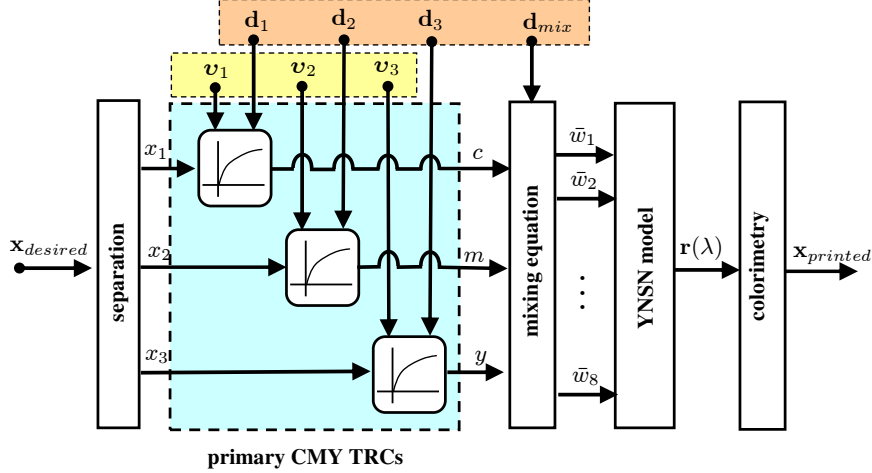


Figure 2.5: The halftone color printing process based on the Yule-Nielsen Spectral Neugebauer Model (YNSN). In this figure,  $\bar{w}_i := w_i + d_{mix,i}$

follows:

$$\mathbf{r}(\lambda) = \left[ \sum_{i=1}^8 (d_{mix,i} + w_i) \mathbf{r}_{P,i}^{1/n_y}(\lambda) \right]^{n_y} \quad (2.14)$$

where  $d_{mix,i} \in \mathbb{R}$  denotes the mixing variations and the weighted average,  $w_i$  can be described by the random mixing equation i.e.

$$\begin{aligned} w_i \in \{ & (1-c)(1-m)(1-y), (1-c)(1-m)y, \\ & (1-c)m(1-y), (1-c)my, \\ & c(1-m)(1-y), c(1-m)y, \\ & cm(1-y), cmy \} \end{aligned}$$

$[c, m, y] \in [0, 1]^3$  represents the fractional areas covered by the CMY colorants respectively. Let the values of the CMY colorants be given by  $\mathbf{x} := [x_1, x_2, x_3]^T \in [0, 1]^3$  respectively. The relationship between  $[c, m, y]$  values and the digital control input of the CMY colorants given by  $[x_1, x_2, x_3]$  values are described by the TRC of each of the primary CMY colorants as given by (2.2). Hence, for each of the C, M and Y TRC, we have a finite number of  $M_1$ ,  $M_2$  and  $M_3$  monochromatic tones respectively with corresponding  $m_1$ ,  $m_2$  and  $m_3$  number of actuators.

By (2.5) and colorimetry (see Appendix A.1), the predicted reflectance,  $\mathbf{r}(\lambda)$  can then be transformed into a three value perceptually uniform device independent colorspace (e.g. CIE  $L^*a^*b^*$  colorspace) that uniquely quantified the printed color for a standard viewing and lighting condition. In (2.14), the parameters  $n_y$  (Yule-Nielsen factor) and  $\mathbf{r}_{P,i}(\lambda)$



(Neugenbauer primary reflectance) that characterized the halftone color print model can be estimated by optimization approaches described in [35, 36, 37]. These parameters are fixed for a given printer.

Hence with this model, the print variations are due to the disturbances on each of the primary TRCs and the mixing variations. Let,

$$\mathbf{d} := [\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_{mix}]^T \in \mathbb{R}^{M_1} \times \mathbb{R}^{M_2} \times \mathbb{R}^{M_3} \times \mathbb{R}^8 \quad (2.15)$$

where  $\mathbf{d}_1$ ,  $\mathbf{d}_2$  and  $\mathbf{d}_3$  are the unknown disturbances of the C, M and Y primary separations.  $\mathbf{d}_{mix} = [d_{mix,1}, \dots, d_{mix,8}]^T$  is the mixing disturbances introduced in (2.14).

Compensation of these disturbances in the xerographic color printing process can be achieved by making appropriate changes to the xerographic actuator settings

$$\mathbf{v} := [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]^T \in \mathbb{R}^{m_1+m_2+m_3} \quad (2.16)$$

where  $\mathbf{v}_1 \in \mathbb{R}^{m_1}$ ,  $\mathbf{v}_2 \in \mathbb{R}^{m_2}$  and  $\mathbf{v}_3 \in \mathbb{R}^{m_3}$  are the actuator settings for the C, M and Y primary separations.

Let,

$$\boldsymbol{\varphi}(\mathbf{v}, \mathbf{d})[\mathbf{x}] = \begin{pmatrix} \varphi_1(\mathbf{v}, \mathbf{d})[\mathbf{x}] \\ \varphi_2(\mathbf{v}, \mathbf{d})[\mathbf{x}] \\ \varphi_3(\mathbf{v}, \mathbf{d})[\mathbf{x}] \end{pmatrix} \in \mathcal{C} \subset \mathbb{R}^3 \quad (2.17)$$

denotes the printed color in  $L^*a^*b^*$  colorspace for input CMY colorants  $\mathbf{x} \in [0, 1]^3$ . Similar to the case of the TRC, the CRC can be adequately approximated by its response at a finite number of colors with a total of  $M_t = M_1 M_2 M_3$  color combinations. For  $M_1$ ,  $M_2$  and  $M_3$  sufficient large, the colors will be closed together, making the volume between close-by colors that are in-differentiable to the human visual system. Thus, the xerographic color

control process can be represented by:

$$\begin{aligned}
 CRC(kT) &:= \boldsymbol{\varphi}(\mathbf{v}, \mathbf{d}) \\
 &= \begin{pmatrix} \varphi_1(\mathbf{v}, \mathbf{d})[\mathbf{x}_1] \\ \vdots \\ \varphi_1(\mathbf{v}, \mathbf{d})[\mathbf{x}_{M_t}] \\ \varphi_2(\mathbf{v}, \mathbf{d})[\mathbf{x}_1] \\ \vdots \\ \varphi_2(\mathbf{v}, \mathbf{d})[\mathbf{x}_{M_t}] \\ \varphi_3(\mathbf{v}, \mathbf{d})[\mathbf{x}_1] \\ \vdots \\ \varphi_3(\mathbf{v}, \mathbf{d})[\mathbf{x}_{M_t}] \end{pmatrix} \in \mathbb{R}^{3M_t} \quad (2.18)
 \end{aligned}$$

The time index  $k$  typically corresponds to the number of photoreceptor belt revolutions with inter-sampling time,  $T$ . As with the case of the TRC, the nonlinear model in (2.18) is not well known.

## Chapter 3

# Sensing of the Temporal-Spatial Signals

In this chapter, the problem of sensing the high-dimensional time-varying tone reproduction curve (TRC)/color reproduction characteristics (CRC) function under the constraint that only a small number of  $n$  tones/colors samples can be taken at each print cycle is described. The time-varying TRC and CRC can be generally described by the  $1 + N$  dimensional temporal-spatial signals i.e. for the case of the time-varying TRC,  $N = 1$  (for the single tone dimension) and for the case of the time-varying CRC,  $N = 3$  (for the color dimensions) as described in section 3.1. The sensing of the temporal-spatial signals is posed in section 3.2 in the framework of the time-sequential sampling approach. The temporal-spatial signal can be recovered from the time-sequentially sensed signal through a reconstruction process via a Kalman filter. This is described in section 3.3. Finally in section 3.4, a simulated case study is conducted to show the effectiveness and related issues with the time-sequential sampling approach. Throughout this chapter, please refer to the notations given in the beginning of this dissertation.

### 3.1 Temporal-Spatial Signals

In general, both the  $TRC(t)$  and  $CRC(t)$  mapping at all times can be represented by a  $1 + N$  dimensional temporal-spatial signal i.e.

$$\omega(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$$

where  $t \in \mathbb{R}$  and  $\mathbf{x} \in \mathcal{S} \subset \mathbb{R}^N$ . We can denote the domain space of both the TRC and CRC mapping by the spatial space so that  $[0, A] \subset \mathbb{R}$  is the true tonal space for the  $TRC(t)$  (e.g. tonal space) sampling problem and  $[0, A_1] \times [0, A_2] \times [0, A_3] \subset \mathbb{R}^3$  is the color space for the  $CRC(t)$  (e.g. CMY colorspace) sampling problem. Hence, in general

$$\mathcal{S} = \prod_{i=1}^N [0, A_i] \subset \mathbb{R}^N \quad (3.1)$$

where  $A_i > 0$  denotes the range of the  $i$ -th spatial domain. Consider  $\mathcal{S}$  to be uniformly discretized to  $M_i > 0$  spatial points at each  $i$ -th spatial dimension. Following (2.2) and (2.18), denote the  $NM_t$ -dimensional time dependent vector of discretized output of the spatial signal by:

$$\bar{\omega}(k) := \begin{pmatrix} \omega_1(kT, \mathbf{x}_1) \\ \vdots \\ \omega_1(kT, \mathbf{x}_{M_t}) \\ \dots \\ \vdots \\ \dots \\ \omega_N(kT, \mathbf{x}_1) \\ \vdots \\ \omega_N(kT, \mathbf{x}_{M_t}) \end{pmatrix} \in \mathbb{R}^{NM_t} \quad (3.2)$$

where  $M_t = M_1$  for the case of the TRC and  $M_t = M_1 M_2 M_3$  for the case of the CRC. In general  $M_t = \text{prod}(\mathbf{M})$  where  $\mathbf{M} = [M_1, \dots, M_N]^T$ . Then, the control process can be given by:

$$\bar{\omega}(k) = \varphi(\mathbf{v}, \mathbf{d}) \in \mathbb{R}^{NM_t} \quad (3.3)$$

where  $\mathbf{v} \in \mathbb{R}^{m_t}$  denotes the control input.  $m_t = m_1$  for the case of the TRC and  $m_t = m_1 + m_2 + m_3$  for the case of the CRC where  $m_i$  corresponds to the number of actuators available for each of the  $i$ -th primary separation.  $\mathbf{d} \in \mathbb{R}^{n_d}$  is the disturbance

**Remark 3.1**

1. A nominal  $\bar{\omega}(k)$  is generally designed so that it can be achieved when no disturbance is present. Let it be  $\bar{\omega}^* := \varphi(\mathbf{v}_0, \mathbf{0})$  where  $\mathbf{v}_0 \in \mathbb{R}^{m_t}$  is the nominal control.
2. It is assumed that  $\varphi(\mathbf{v}, \mathbf{d})$  is differentiable and defined for all values of  $\mathbf{v}$  and  $\mathbf{d}$ .

### 3.1.1 Linear static plant with uncertainty

The nonlinear model given by (3.3) is not well known. To account for the effect of disturbances and the actuator settings, a set of uncertain static linear models are used instead [3]. This is possible because of the static nature of the xerographic printing process.

Let  $\bar{\mathbf{v}} := \mathbf{v} - \mathbf{v}_0$  be the deviation from the nominal control. Define the TRC/CRC error to be:

$$\bar{\mathbf{e}}(k) = \bar{\boldsymbol{\omega}}(k) - \bar{\boldsymbol{\omega}}^* \quad (3.4)$$

$$= \boldsymbol{\varphi}(\mathbf{v}, \mathbf{d}) - \boldsymbol{\varphi}(\mathbf{v}_0, \mathbf{0}) \in \mathbb{R}^{NM_t} \quad (3.5)$$

By mean value theorem we have:

$$\boldsymbol{\varphi}(\mathbf{v}, \mathbf{d}) = \boldsymbol{\varphi}(\mathbf{v}_0, \mathbf{d}) + \hat{\boldsymbol{\phi}} \cdot \bar{\mathbf{v}} + \bar{\Delta}(\mathbf{v}, \mathbf{d})\bar{\mathbf{v}}$$

where  $\hat{\boldsymbol{\phi}}$  gives the nominal sensitivity matrix. Assume here that  $\bar{\Delta}(\mathbf{v}, \mathbf{d})$  fulfills the following condition for all times.

$$\text{range}(\bar{\Delta}(\mathbf{v}, \mathbf{d})) \subset \text{range}(\hat{\boldsymbol{\phi}}) \quad (3.6)$$

Hence, we can define, for each time step  $k$ , the uncertainty  $\Delta(k) \in \mathbb{R}^{m_t \times m_t}$  and weighting  $\mathbf{W}_v \in \mathbb{R}^{m_t \times m_t}$  so that:

$$\bar{\Delta}(\mathbf{v}, \mathbf{d}) = \hat{\boldsymbol{\phi}}\Delta(k)\mathbf{W}_v$$

Let

$$\boldsymbol{\varphi}(\mathbf{v}_0, \mathbf{d}) - \boldsymbol{\varphi}(\mathbf{v}_0, \mathbf{0}) =: \bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$$

Hence from (3.5), the linear uncertain model is given by:

$$\bar{\mathbf{e}}(k) = \hat{\boldsymbol{\phi}}(\mathbf{I} + \Delta(k)\mathbf{W}_v)\bar{\mathbf{v}}(k) + \bar{\mathbf{d}}(k) \quad (3.7)$$

#### Remark 3.2

1. Note that if the condition given by (3.6) is fulfilled, no approximation is made with the linear uncertain model given in (3.7).
2. When assumption (3.6) is not satisfied, the discrepancy between the linear uncertain model in (3.7) and the nonlinear plant in (3.5) is given by the portion of  $\bar{\Delta}(\mathbf{v}, \mathbf{d})\bar{\mathbf{v}}$  which is not in the range space of  $\hat{\boldsymbol{\phi}}$ . Since  $\bar{\Delta}(\mathbf{v}, \mathbf{d})$  will be small when  $\mathbf{d} \approx \mathbf{0}$ , the discrepancy corresponds to a second-order term with respect to small  $\bar{\mathbf{v}}$  and  $\mathbf{d}$ .

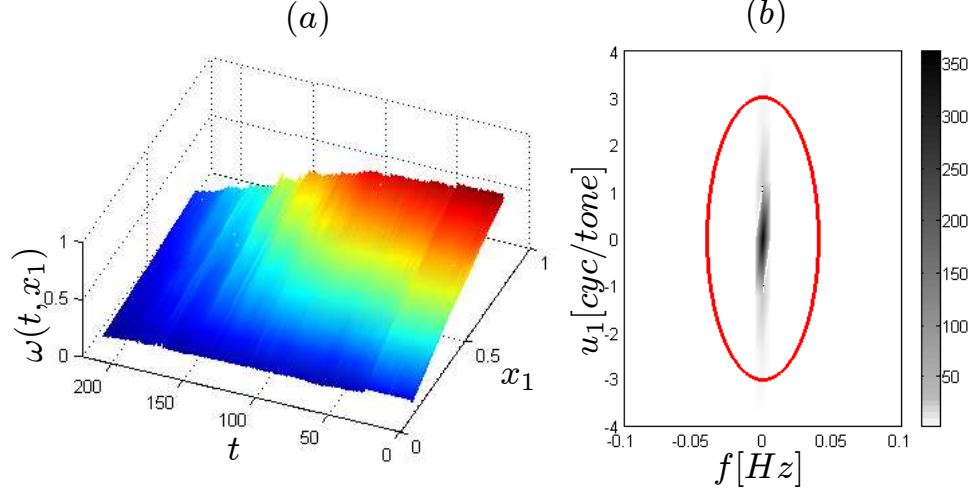


Figure 3.1: (a) The time-varying TRC signal,  $\omega(t, x_1)$  from a legacy printer at nominal actuator inputs  $\mathbf{v}_0$ ; (b) The frequency content of the time-varying TRC with the assumed ellipsoidal spectral content,  $\Theta(W, U_1)$

### 3.1.2 Disturbance model

The TRC/CRC disturbance can be analyzed by looking at the frequency content of the time-varying TRC/CRC (i.e. the temporal-spatial signal) at nominal actuator inputs  $\mathbf{v}_0$ . Figure 3.1(a) shows the time-varying TRC,  $\omega(t, x_1)$  from a legacy printer at nominal actuator inputs  $\mathbf{v}_0$ . Let  $\Omega(f, \mathbf{u})$  be the power-spectral density of  $\omega(t, \mathbf{x})$  at nominal actuator input  $\mathbf{v}_0$ , where  $f$  is the temporal frequency and  $\mathbf{u} = [u_1, \dots, u_N]$  is the spatial frequency. Since the spatial domain  $\mathcal{S}$  is finite,  $\Omega(f, \mathbf{u})$  cannot be truly band-limited. In this dissertation, following [13], it is assumed that  $\Omega(f, \mathbf{u})$  is essentially band-limited to a closed support region,  $\Theta \subset \mathbb{R} \times \mathbb{R}^N$  in the sense that

$$\int_{\mathbb{R} \times \mathbb{R}^N / \Theta} \Omega(f, \mathbf{u}) df d\mathbf{u} \leq \epsilon^2 \quad (3.8)$$

so that all but  $\epsilon^2$  of the spectral energy is concentrated in  $\Theta$ . We further assume that this compact spectral support region is given by an ellipse with principal axes aligned with the coordinate axes i.e.

$$\Theta(W, U_1) = \left\{ (f, u_1) : \frac{f^2}{W^2} + \frac{u_1^2}{U_1^2} \leq 1 \right\} \subset \mathbb{R} \times \mathbb{R} \quad (3.9)$$

This is illustrated in Figure 3.1(b). The ellipsoidal body  $\Theta$  is assumed because the signal of interest to us i.e. the time-varying TRC has significant frequency content that can be grouped in an ellipsoid region (see Figure 3.1(b) with the high frequency variations more

visibly demonstrated in Figure 3.2(d)). The frequency content is obtained from a two-dimensional Fourier analysis on a track of experimental time-varying TRC data. The data is obtained from measurement of 9 tones variations patches by running a legacy printer at open loop with nominal actuator settings for an extended time-period.

Similarly, this is assumed to be true for the case of the time-varying CRC. In general, we assume that the compact spectral support region is given by ellipsoids with principle axes aligned with the coordinate axes, i.e.

$$\Theta(W, \mathbf{U}) = \left\{ (f, \mathbf{u}) : \frac{f^2}{W^2} + \sum_{i=1}^N \frac{u_i^2}{U_i^2} \leq 1 \right\} \subset \mathbb{R} \times \mathbb{R}^N \quad (3.10)$$

where  $W$  is the length along the temporal frequency direction and  $\mathbf{U} = [U_1, \dots, U_N]$  are the lengths along the spatial frequency direction. Note that the restriction that the ellipsoidal body always aligns with the input color coordinates are not so important here since we can always re-coordinate such that the form given by (3.10) is achieved. The parameterizations of the compact spectral support  $\Theta$  will henceforth be given by known values of  $(W, \mathbf{U})$ . To determine the compact spectral support  $\Theta(W, \mathbf{U})$ , the procedure given in Appendix B.2 is used. In this procedure a threshold value,  $\delta_t \in \mathbb{R}$  is selected such that we obtained the significant power spectral densities in  $\Omega(f, \mathbf{u})$  that are tightly wrapped around by an ellipsoidal region.  $\delta_t$  is selected to be small here to include all temporal-spatial frequency contents that are above the noise level<sup>1</sup>

The Fourier analysis of the temporal-spatial signal (in this case the time-varying TRC) enables determination of the region where all the frequency content of the signal lies. This is critical in the designing effective time-sequential sampling strategy that minimizes aliasing. In addition, motivated by the experimental time-varying TRC data, this class of temporal-spatial (tonal) signal exhibits the following unique characteristics:

1. Higher output tone variations can be observed at high input tones as shown in Figure 3.2(a). These variations are slow varying lying within a small narrow temporal frequency band as shown in Figure 3.2(b). We called these variations the *low-frequency band* variations.
2. Additionally there are also relatively small magnitude output tone variations that are *not* input tone dependent as shown in Figure 3.2(c). In this data, these variations occurs at high-temporal frequency as shown in Figure 3.2(d). We called these variations

---

<sup>1</sup>The noise level is determined by frequency analysis of the variations data from repeated measurements of the same print.

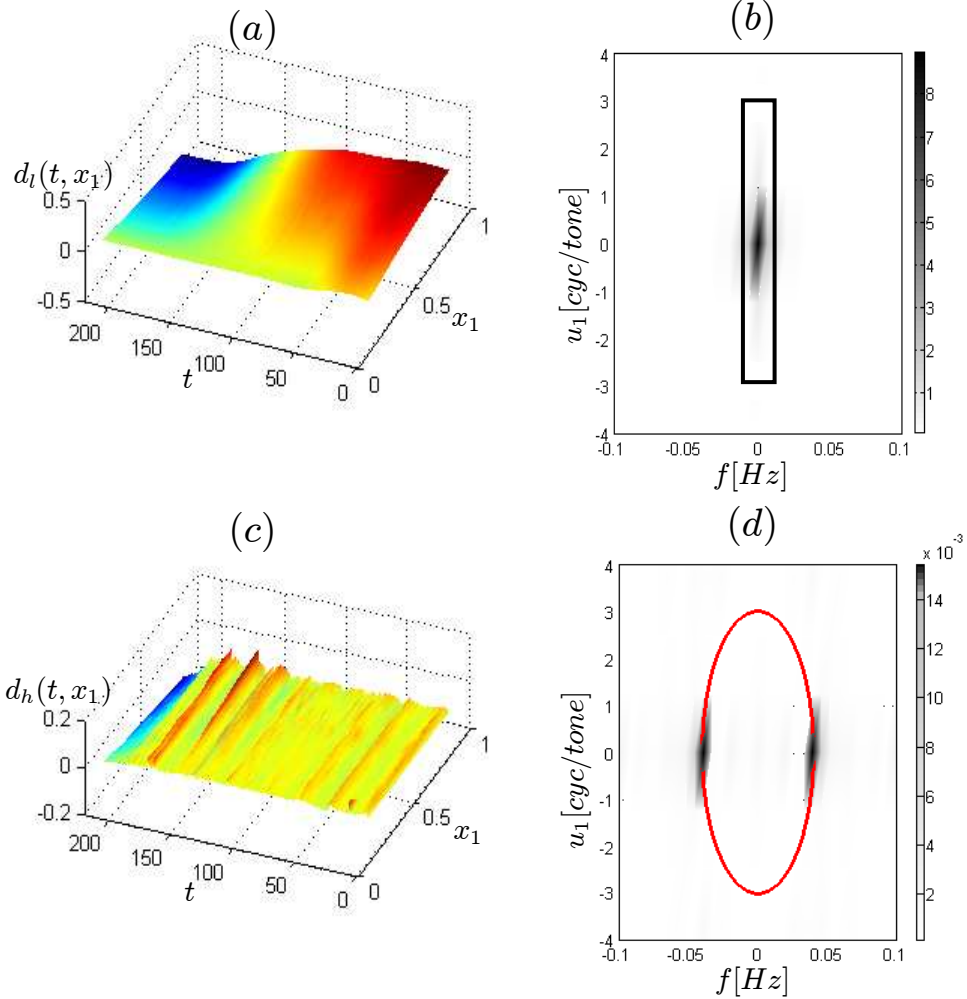


Figure 3.2: (a) The low-frequency band TRC variations from the nominal TRC given by  $d_l(t, x)$  at nominal actuator inputs  $\mathbf{v}_0$ ; (b) Corresponding power spectral density of  $d_l(t, x)$ . The low-frequency band is given by the rectangular box region; (c) The high-frequency band TRC variations from the nominal TRC given by  $d_h(t, x)$  at nominal actuator inputs  $\mathbf{v}_0$ ; (d) Corresponding power spectral density of  $d_h(t, x)$ . The high-frequency band can be characterized to lie in an ellipsoidal region for a given threshold value setting for determination of the significant power spectral density (see Appendix B.2).



the *high-frequency band* variations.

We assumed here that these unique characteristics as observed in the time-varying TRC can be similarly extended to the case of the CRC. The different bands of variations fit well to the context of the xerographic color printing systems as the time scales of the factors influencing these variations of these systems (environmental, materials, media etc.), vary considerably. Consideration of these unique characteristics are important in designing an effective causal reconstruction filter of the time-sequentially sampled signal based of Kalman filter (see section 3.3) and for a realistic simulation of the disturbances of the print process.

The variations of the time-varying TRC/CRC is introduced into (3.7), by the disturbances,  $\bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$  variable ( $N = 1$  for the case of the TRC and  $N = 3$  for the case of the CRC). To exhibit the spectral contents and unique characteristics of the time-varying TRC/CRC, the disturbances,  $\bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$  is modeled as follows:

$$\bar{\mathbf{d}}(k) = \mathbf{F}_\ell \cdot \mathbf{x}_\ell(k) + \mathbf{F}_h \cdot \mathbf{x}_h(k) \quad (3.11)$$

where  $k \in \mathbb{Z}^+$  is the time index. This model is known as the two-frequency bands model.

$\mathbf{F}_\ell \cdot \mathbf{x}_\ell(k)$  gives the low-frequency band variations of the TRC/CRC. In this formulation:

- $\mathbf{F}_\ell \in \mathbb{R}^{NM_t \times NM_t}$  is the spatial basis matrix derived from a representative TRC/CRC data track,  $\bar{\omega}(k) \in \mathbb{R}^{N\hat{M}_t}$  collected for  $k = 0$  to  $K - 1$ ,  $K - 1 > N\hat{M}_t$  at nominal actuator inputs  $\mathbf{v}_0$ . From (3.7), at nominal  $\bar{\mathbf{d}}(k) = \bar{\mathbf{e}}(k)$  since  $\bar{\mathbf{v}}_0 = \mathbf{0}$ . Let  $\mathbf{\Gamma} = [\bar{\omega}(0) - \bar{\omega}^*, \dots, \bar{\omega}(K - 1) - \bar{\omega}^*] \in \mathbb{R}^{N\hat{M}_t \times K}$ . By the singular value decomposition (svd) we have

$$\mathbf{\Gamma} = \mathbf{\Lambda} [\mathbf{\Sigma}_1 \quad \mathbf{\Sigma}_2] \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}$$

The left singular matrix  $\mathbf{\Lambda} \in \mathbb{R}^{N\hat{M}_t \times N\hat{M}_t}$  spans the column space of  $\mathbf{\Gamma}$  i.e the spatial subspace.  $\mathbf{\Sigma}_1 \in \mathbb{R}^{N\hat{M}_t \times N\hat{M}_t}$  and  $\mathbf{\Sigma}_2 \in \mathbb{R}^{N\hat{M}_t \times (K - N\hat{M}_t)}$  give the diagonal singular matrices.  $\mathbf{V}_1^T \in \mathbb{R}^{N\hat{M}_t \times K}$  and  $\mathbf{V}_2^T \in \mathbb{R}^{(K - N\hat{M}_t) \times K}$  give the right singular matrices that spans the row space of  $\mathbf{\Gamma}$  i.e the temporal subspace. Hence, the spatial basis matrix  $\mathbf{F}_\ell$  of the disturbance is given by:

$$\mathbf{F}_\ell = \mathbf{\Lambda} \mathbf{\Sigma}_1$$

With this spatial basis matrix, the different level of output tone/color variations at different input tones/colors as clearly shown in Figure 3.2(a) for the case of the TRC can be captured.

- $\mathbf{x}_\ell(k) \in \mathbb{R}^{NM_t}$  captures the time variations of the TRC/CRC. It is modeled as follows:

$$\begin{aligned}\mathbf{x}_1(k+1) &= \mathbf{A}_1\mathbf{x}_1(k) + \mathbf{B}_1\mathbf{w}_1(k) \\ \mathbf{x}_\ell(k) &= \mathbf{C}_1\mathbf{x}_1(k) + \mathbf{D}_1\mathbf{w}_1(k)\end{aligned}\tag{3.12}$$

where  $\mathbf{w}_1(k)$  is a white process noise with covariance  $\mathbf{R}_1$ . The matrix

$$\begin{aligned}\mathbf{A}_1 &= \begin{pmatrix} \mathbf{A}_{1,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{1,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{1,N} \end{pmatrix} \in \mathbb{R}^{NM_t n_{f_1} \times NM_t n_{f_1}} \\ \mathbf{B}_1 &= \begin{pmatrix} \mathbf{B}_{1,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{1,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}_{1,N} \end{pmatrix} \in \mathbb{R}^{NM_t n_{f_1} \times NM_t} \\ \mathbf{C}_1 &= \begin{pmatrix} \mathbf{C}_{1,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{1,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}_{1,N} \end{pmatrix} \in \mathbb{R}^{NM_t \times NM_t n_{f_1}} \\ \mathbf{D}_1 &= \begin{pmatrix} \mathbf{D}_{1,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{1,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}_{1,N} \end{pmatrix} \in \mathbb{R}^{NM_t \times NM_t}\end{aligned}$$

where  $(\mathbf{A}_{1,j}, \mathbf{B}_{1,j}, \mathbf{C}_{1,j}, \mathbf{D}_{1,j})$  for  $j = 1, \dots, N$  are given by a  $M_t$  bank of  $n_{f_1}$ -order low-pass filters that filter each spatial channel with a constant temporal cutoff frequencies,  $cf_j$  [Hz].

$\mathbf{F}_h \cdot \mathbf{x}_h(k)$  gives the high-frequency band variations with temporal-spatial spectral content that lies in an ellipsoidal region as shown in Figure 3.2(d). To shape the temporal-spatial spectral content, a pink-noise disturbance dynamics as previously given in [23] is used. In this modeling approach,

- $\mathbf{F}_h \in \mathbb{R}^{NM_t \times NM_t}$  is a matrix of Fourier basis matrix (see derivation of this basis function for the case of  $N = 1$  (TRC) and  $N = 3$  (CRC) in Appendix B.1) derived from the inverse discrete Fourier transform formulation.  $\mathbf{F}_h$  differs from the spatial basis matrix,  $\mathbf{F}_\ell$  in that  $\mathbf{F}_h$  basis vectors treat the variations uniformly throughout the

spatial domain (i.e. spatial uniformity) whereas  $\mathbf{F}_\ell$  basis vectors treat the variations to be higher at high values of the spatial domain (i.e. high input tones/colors).

- $\mathbf{x}_h(k)$  is the vector of Fourier coefficients. The temporal frequency content at each spatial frequency channel can be shaped through  $\mathbf{x}_h(k)$ . Hence,  $\mathbf{x}_h(k)$  is modeled by the following dynamics:

$$\begin{aligned}\mathbf{x}_2(k+1) &= \mathbf{A}_2\mathbf{x}_2(k) + \mathbf{B}_2\mathbf{w}_2(k) \\ \mathbf{x}_h(k) &= \mathbf{C}_2\mathbf{x}_2(k) + \mathbf{D}_2\mathbf{w}_2(k)\end{aligned}\tag{3.13}$$

where  $\mathbf{w}_2(k)$  is a white process noise with covariance  $\mathbf{R}_2$ . The matrix

$$\begin{aligned}\mathbf{A}_2 &= \begin{pmatrix} \mathbf{A}_{2,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{2,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{A}_{2,N} \end{pmatrix} \in \mathbb{R}^{NM_t n_{f_2} \times NM_t n_{f_2}} \\ \mathbf{B}_2 &= \begin{pmatrix} \mathbf{B}_{2,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{2,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{B}_{2,N} \end{pmatrix} \in \mathbb{R}^{NM_t n_{f_2} \times NM_t} \\ \mathbf{C}_2 &= \begin{pmatrix} \mathbf{C}_{2,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{2,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{C}_{2,N} \end{pmatrix} \in \mathbb{R}^{NM_t \times NM_t n_{f_2}} \\ \mathbf{D}_2 &= \begin{pmatrix} \mathbf{D}_{2,1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{2,2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}_{2,N} \end{pmatrix} \in \mathbb{R}^{NM_t \times NM_t}\end{aligned}$$

where  $(\mathbf{A}_{2,j}, \mathbf{B}_{2,j}, \mathbf{C}_{2,j}, \mathbf{D}_{2,j})$  for  $j = 1, \dots, N$  are given by a  $M_t$  bank of  $n_{f_2}$ -order low-pass filters that filter each spatial channel (each Fourier coefficient) with temporal cutoff frequencies corresponding to an ellipsoids as given by (3.10). This gives an approximation of a compact ellipsoidal temporal-spatial spectral support.

## 3.2 Time-Sequential Sampling

In this dissertation, time-sequential sampling refers to a periodic information operator  $TS(n) : \boldsymbol{\omega}(t, \mathbf{x}) \rightarrow \boldsymbol{\omega}_s(t), \boldsymbol{\omega}_s(t) \in \mathbb{R}^n$  that samples  $n$  pieces of spatial information,  $\boldsymbol{\omega}(t, \cdot)$  at each sampling instance  $kT$ , where  $k \in \mathbb{Z}^+$  and  $T$  is the inter-sampling time.

### 3.2.1 Framework of time-sequential sampling

Sensing of the  $\boldsymbol{\omega}(t, \mathbf{x})$  at time instant  $t = kT$  is achieved by printing and measuring  $n$  tones in the form of small test patches. Typically in current printers, the same set of  $n$  tones is printed at each  $k$  and  $n$  will be determined by the number of available sensors, as well as the productivity and materials cost of printing the test patches. This dissertation proposes to print  $n$  *different* tones at different time  $k$ . We called such sampling strategy a time-sequential sampling in that the sampled tones/colors is a function of time. Which tones are printed at what times are determined by the periodic  $TS(n)$  sampling.

To develop the time-sequential sampling framework, the tone/color domain of the temporal-spatial signal,  $\boldsymbol{\omega}(t, \mathbf{x}), \mathbf{x} \in \mathcal{S}$  is given by (3.1) i.e.:

$$\mathcal{S} = \prod_{i=1}^N [0, A_i) \subset \mathbb{R}^N$$

Let the set of possible sampling locations be given by a rectangular grid:

$$\mathcal{G} = \{(t, \mathbf{x}) \mid t = kT, \mathbf{x} = \mathbf{X} \cdot * \boldsymbol{\ell}, \forall (k, \boldsymbol{\ell}) \in \mathbb{Z}^{1+N}\} \quad (3.14)$$

where  $T$  gives the inter-sampling time and  $\mathbf{X} = \mathbf{A} \cdot / \mathbf{M}$  gives the  $N$ -dimensional spatial resolution.  $\mathbf{A} = [A_1, \dots, A_N] \in \mathbb{R}^{+N}$  is the  $N$ -dimensional spatial range of  $\mathcal{S}$  and  $\mathbf{M} = [M_1, \dots, M_N] \in \mathbb{Z}^{+N}$  is the  $N$ -dimensional number of points that uniformly discretize the spatial range. These parameterizations are shown in Figure 3.3 for the case of  $N = 2$  spatial dimension. In (3.14),  $\boldsymbol{\ell} \in \mathbb{Z}^N$  gives the  $N$ -tuples integer indices that specify the location of the grid points. Let  $\mathcal{E}_\alpha$  be the set of *grid indices* such that:

$$\mathcal{E}_\alpha = \{\boldsymbol{\ell} \mid \mathbf{X} \cdot * \boldsymbol{\ell} \in \mathcal{S}\} \quad (3.15)$$

Because  $\mathcal{S}$  is a rectangular domain,  $\mathcal{E}_\alpha$  can be enumerated as the  $N$ -tuples integer permu-



rewritten as:

$$\bar{\omega}(k) = \begin{pmatrix} \omega_1(kT, \mathbf{X} \cdot * \ell_1) \\ \vdots \\ \omega_1(kT, \mathbf{X} \cdot * \ell_{M_t}) \\ \dots \\ \vdots \\ \dots \\ \omega_N(kT, \mathbf{X} \cdot * \ell_1) \\ \vdots \\ \omega_N(kT, \mathbf{X} \cdot * \ell_{M_t}) \end{pmatrix} \in \mathbb{R}^{NM_t}$$

where  $\ell_p, p = 1, \dots, M_t$  gives the  $p$ -th enumerated member of the set  $\mathcal{E}_\alpha$ .

An  $n$ -point time-sequential sampling,  $TS(n)$  means that  $n$  spatial points are sampled at temporal-spatial locations  $[(kT, \mathbf{X} \cdot * \alpha_1(k)), \dots, (kT, \mathbf{X} \cdot * \alpha_n(k))]$  on the grid,  $\mathcal{G}$ . At each spatial point,

$$\alpha_i(k) \in \mathcal{E}_\alpha \quad (3.17)$$

for  $i \in \{1, \dots, n\}$  gives the time-dependent index order sequence in which the  $i$ -th spatial samples are taken. From (3.16),  $\mathcal{E}_\alpha$  is finite with cardinality  $M_t = \text{prod}(\mathbf{M})$ , hence for all  $i \in \{1, \dots, n\}$ ,  $\alpha_i(k)$  takes unique integer vectors (i.e. unique sampling points in the grid  $\mathcal{G}$  with respect to the spatial domain  $\mathcal{S}$ ) for  $k = 0, \dots, M_p - 1$  where  $M_p$  is the periodicity of the sequence. Outside of this time indices, the sequence repeats itself periodically such that  $\alpha_i(k) = \alpha_i(k + M_p)$ . Let

$$\alpha(k) = [\alpha_1(k), \dots, \alpha_n(k)] \quad (3.18)$$

Hence, the  $TS(n)$  sampling is defined by the quadruple of  $(\alpha(k), T, \mathbf{A}, \mathbf{M})$ .

It is useful here to look at the  $TS(n)$  sampling as a product of an indicator matrix sequence that depends on the  $TS(n)$  sampling sequence,  $\alpha(k)$  and the time-dependent vector of discretized output of the spatial signal given by (3.2). First notice that since  $\mathcal{E}_\alpha$  is finite it is enumerable. Let

$$f_{enum} : \mathcal{E}_\alpha \rightarrow \{1, 2, \dots, M_t\} \quad (3.19)$$

be a function that enumerates all elements of the set  $\mathcal{E}_\alpha$  to an integer number that starts with 1 and ends with  $M_t$ . Hence,  $\alpha(k)$  defines an indicator matrix sequence  $\tilde{\mathbf{C}}_\alpha(k) \in \mathbb{R}^{n \times M_t}$  such that the  $(p, q)$  element of  $\tilde{\mathbf{C}}_\alpha(k)$  is 1 when  $q = f_{enum}(\alpha_p(k))$  and it is 0 otherwise, for

all  $p = 1, \dots, n$ . The  $TS(n)$  sampled temporal-spatial signal is therefore given by:

$$\bar{\omega}_s(k) = \underbrace{\begin{pmatrix} \tilde{\mathbf{C}}_{\alpha}(k) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}_{\alpha}(k) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \tilde{\mathbf{C}}_{\alpha}(k) \end{pmatrix}}_{\mathbf{C}_{\alpha}(k) \in \mathbb{R}^{nN \times nM_t}} \bar{\omega}(k) \in \mathbb{R}^{nN}. \quad (3.20)$$

An example of the  $TS(n)$  sampling is shown in Figure 3.4 for  $n = 1$  and for  $1 + N = 2$  dimensional temporal-tonal signal (i.e. time-varying TRC). Here, the lexicographic and bit-reversed  $TS(1)$  sampling are shown, where

1. **Lexicographic:** Here  $\alpha_1(k) := k \bmod M_1$  so that the order of sampling is according to its index.

$$\alpha_1(k) : 0, 1, \dots, M_1 - 1, 0, 1, \dots, \quad (3.21)$$

2. **Bit reversed:** Here  $\alpha_1(k)$  is given by first representing the index  $k$  as a binary number, then reversing the order of the significant bits. Thus for  $M_1 = 16 = 2^4$ ,

$$\alpha_1(k) : 0, 8, 4, 12, 2, 10, 6, 14, 1, 9, 5, 13, 3, 11, 7, 15, 0, 8, \dots \quad (3.22)$$

It is apparent that the lexicographical sequence does not cover the temporal-tonal domain as uniformly as the bit-reversed sequence. Intuitively the bit-reversed  $TS(1)$  sampling would be more effectively in sampling the temporal-tonal signal compared to the lexicographic  $TS(1)$  sampling. This is analyzed in the next section.

### 3.2.2 Spectral content of time-sequentially sampled signal

To analyze the power spectral density of time-sequentially sampled signal for the lexicographic and bit-reversed sequences,  $TS(1)$  sampling (given by (3.21) and (3.22) respectively for the case of the TRC), note that  $\alpha(k) = \alpha_1(k)$ . We assume that the  $1 + N$  dimensional temporal-spatial signal is a *wide-sense stationary* random process. With this assumption the normalized power spectral density  $\Omega_s(f, \mathbf{u})$  of the  $TS(1)$  sampled signal  $\omega(kT, \mathbf{X} \cdot * \alpha_1(k))$  is given as follows [11, 12, 38]:

$$\Omega_s(f, \mathbf{u}) = \sum_{m, \mathbf{p}} |Q_{m\mathbf{p}}|^2 \Omega(f - m/B, \mathbf{u} - \mathbf{p} \cdot / \mathbf{A}) \quad (3.23)$$

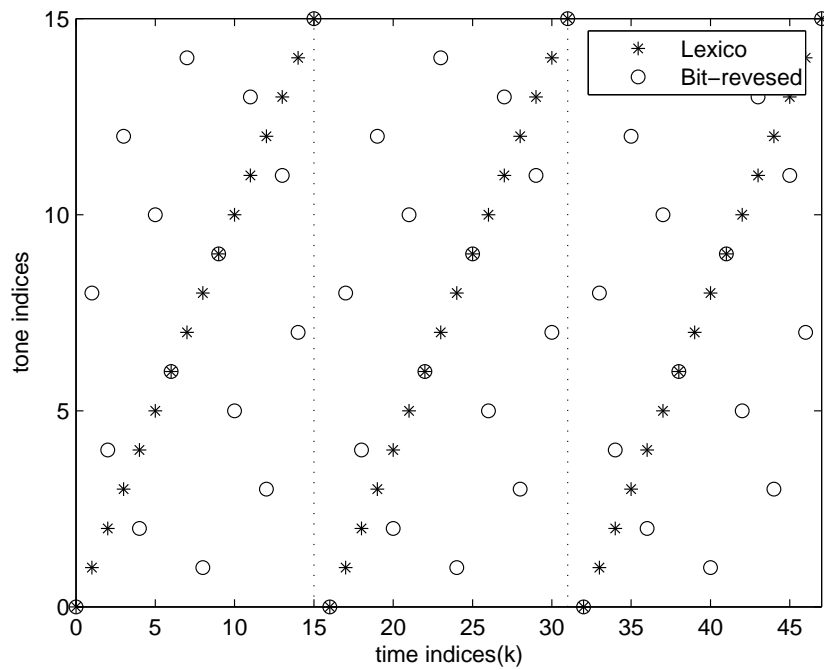


Figure 3.4: Lexicographical and bit-reversed time-sequential sampling sequence with  $M_1 = 16$  tones.



$$Q_{m\mathbf{p}} = \frac{1}{M_t} \sum_{k=0}^{M_t-1} \exp^{-i2\pi[mk/M_t + \mathbf{p} \cdot * (\boldsymbol{\alpha}_1(k) \cdot \mathbf{M})^T]} \quad (3.24)$$

where  $(m, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^N$  and  $B = M_t T$  is the periodicity of the sampling sequence. Equation (3.23) shows that the  $TS(1)$  sampled signal spectral,  $\Omega_s(f, \mathbf{u})$  is given by the  $Q_{m\mathbf{p}}$  weighted duplicates of the baseband spectra,  $\Omega(f, \mathbf{u})$  at  $(m/B, \mathbf{p} \cdot / \mathbf{A})$  position in the temporal-spatial frequency domain. Hence to observe aliasing or overlapping of the weighted baseband, it is critical here to look at the  $Q_{m\mathbf{p}}$  weights. For any  $TS(1)$  sampling sequence,  $\boldsymbol{\alpha}_1(k)$ :

$$Q_{00} = 1 \quad (3.25)$$

$$Q_{0\mathbf{p}} = \boldsymbol{\delta}_{\mathbf{p} \cdot \text{mod } \mathbf{M}} \quad (3.26)$$

$$Q_{m0} = \delta_{m \text{ mod } M_t} \quad (3.27)$$

$$Q_{m+c*M_t, \mathbf{p}+\mathbf{d}*\mathbf{M}} = Q_{m\mathbf{p}}, \text{ for integer } c \text{ and } \mathbf{d} \quad (3.28)$$

$$Q_{m\mathbf{p}} = Q_{-m, -\mathbf{p}} \quad (3.29)$$

The functions  $\delta_{(\cdot)}$  and  $\boldsymbol{\delta}_{(\cdot)}$  are the one and multidimensional Kronecker delta functions respectively. From (3.26), at  $m = 0$  hyperplane, the non-zero  $Q_{0\mathbf{p}}$  that are closest to the origin (location of the baseband) are located  $\mathbf{p} = (M_1, 0, \dots, 0), (0, M_2, \dots, 0), \dots, (0, 0, \dots, M_N)$ . Hence, from (3.23), the original frequency content at zero temporal hyperplane is replicated at  $\mathbf{u} = (M_1/A_1, 0, \dots, 0), (0, M_2/A_2, \dots, 0), \dots, (0, 0, \dots, M_N/A_N)$ . The maximum spatial frequencies are given by  $\mathbf{U} = [U_1, U_2, \dots, U_N]$ , hence no replication will overlap the baseband at the zero temporal hyperplane if

$$U_1 \leq M_1/2A_1, U_2 \leq M_2/2A_2, \dots, U_N \leq M_N/2A_N \quad (3.30)$$

or compactly we write  $\mathbf{U} \leq \mathbf{M} \cdot / 2\mathbf{A}$ . This is known as the *spatial Nyquist* condition. This condition is illustrated for the case of the temporal-spatial signal with  $N = 1$  (e.g. time-varying TRC) in Figure 3.5. Thus, to avoid aliasing, we need to pick  $\mathbf{M} \geq 2\mathbf{A} \cdot * \mathbf{U}$  for known values of spatial ranges  $\mathbf{A}$  and maximum spatial frequencies,  $\mathbf{U}$ .

In addition, from (3.23) the replications of the original frequency content,  $\Omega(f, \mathbf{u})$  in the  $p = \pm 1$  hyperplanes are located *at least*  $1/B$ ,  $B = M_t T$  from the origin. Since the maximum temporal frequency is given by  $W$ , the temporal Nyquist condition is given by:

$$W \leq 1/2B \quad (3.31)$$

Notice that the temporal Nyquist rate  $1/B$  can be less than  $2W$  if  $Q_{m\mathbf{p}} = 0$  for values of  $(m, \mathbf{p})$  other than those indicated in (3.26). This can be achieved by selecting the “right”

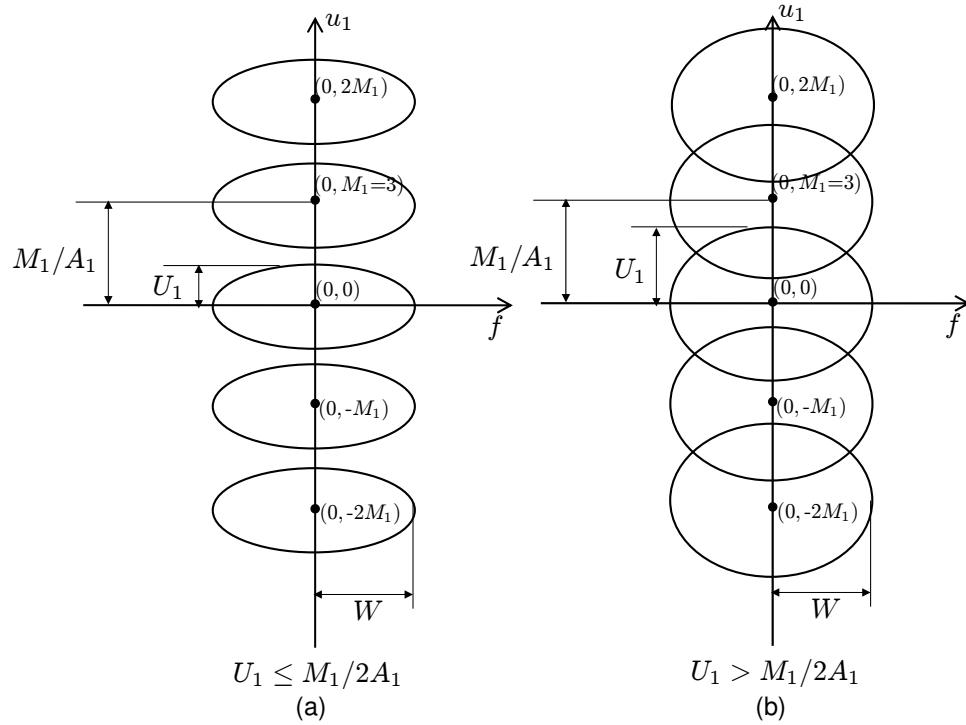


Figure 3.5: (a) Spatial Nyquist condition is fulfilled making the replications of the original frequency content,  $\Omega(f, \mathbf{u})$  (given by the ellipses) in the  $p = 0$  hyperplanes that do not overlap the original frequency content (i.e. the baseband); (b) Spatial Nyquist condition is *not* fulfilled making the replications of the original frequency content,  $\Omega(f, \mathbf{u})$  (given by the ellipses) in the  $p = 0$  hyperplane that overlap the baseband

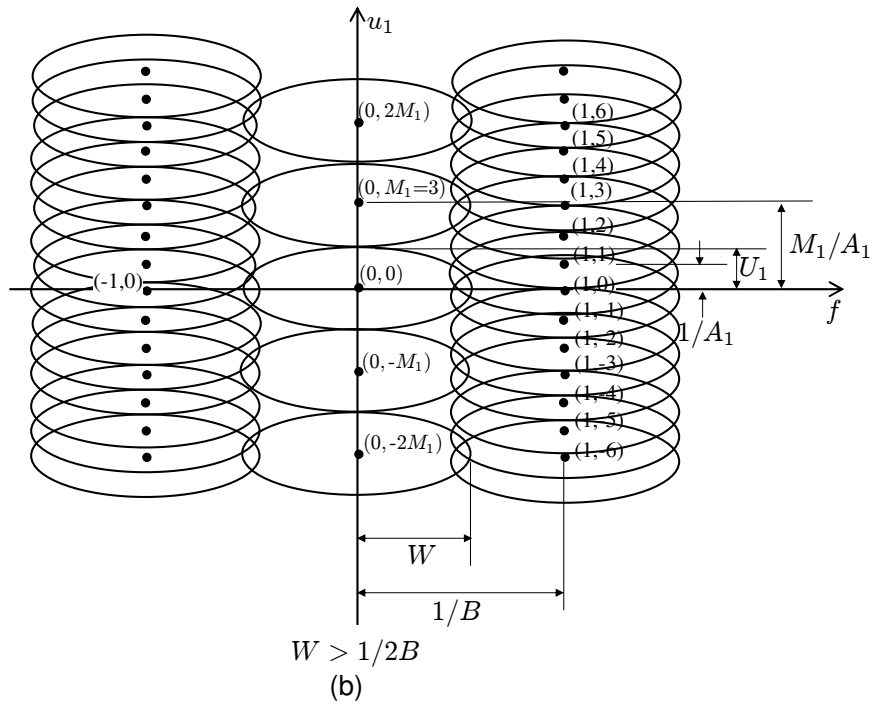
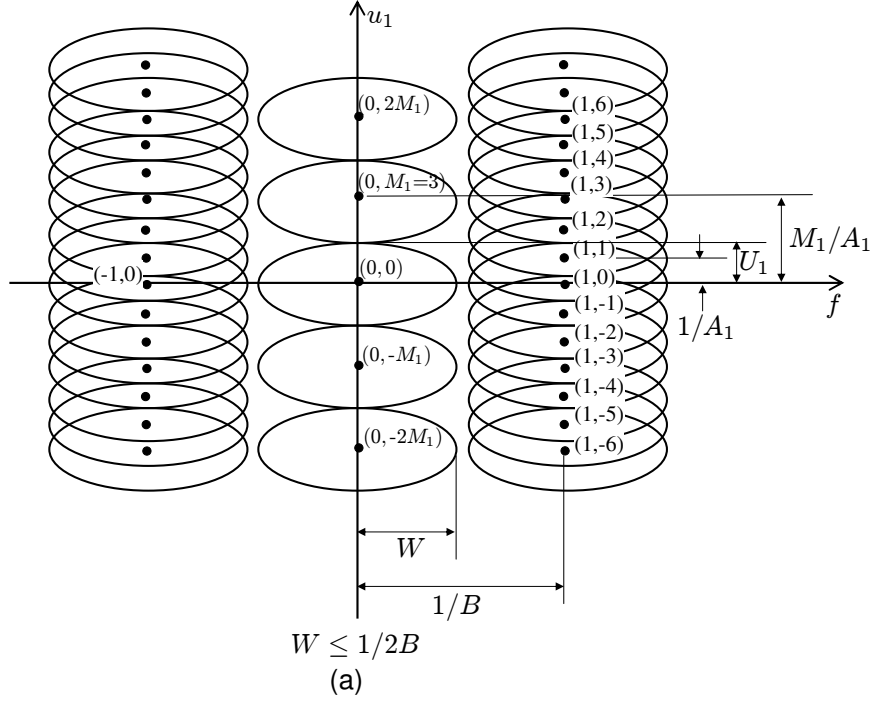


Figure 3.6: (a) Temporal Nyquist condition is fulfilled making the replications of the original frequency content,  $\Omega(f, \mathbf{u})$  (given by the ellipses) in the  $p = \pm 1$  hyperplanes that do not overlap the original frequency content (i.e. the baseband); (b) Temporal Nyquist condition is *not* fulfilled making some of the replications of the original frequency content,  $\Omega(f, \mathbf{u})$  (given by the ellipses) in the  $p = \pm 1$  hyperplanes that overlap the baseband

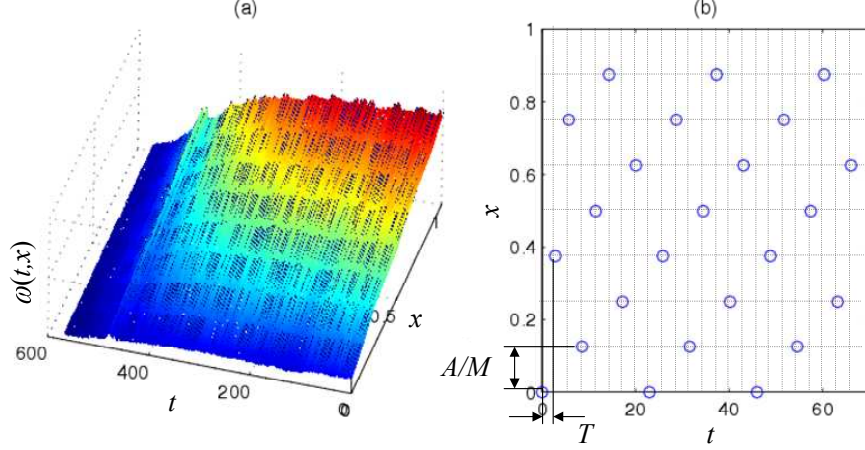


Figure 3.7: (a) The time-varying TRC signal,  $\omega(t, x_1)[1 + N = 2]$  from actual experimental data; (b)  $TS(1)$  sampling of  $\omega(t, x_1)$

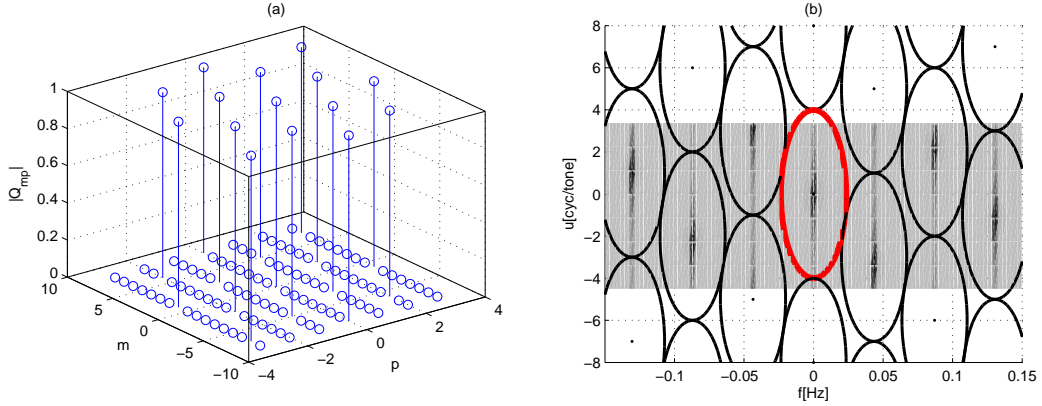


Figure 3.8: (a) Aliasing weights  $|Q_{mp_1}|$  for the  $TS(1)$  sampling shown in Figure 3.7(b); (b) The frequency content  $TS(1)$  sampling of time-varying TRC,  $\Omega_s(f, u_1)$

$TS(1)$  sampling sequence,  $\alpha_1(k)$  that gives the best tradeoff of spatial frequencies for temporal frequency.

As an example, consider the  $TS(1)$  sampling for  $1 + N = 2$  temporal-spatial signal as shown in Figure 3.7(b), the  $|Q_{mp_1}|$  weight is shown in Figure 3.8(a). From (3.23),  $\Omega_s(f, u_1)$  is made up of weighted replication of  $\Omega(f, u_1)$  (see Figure 3.8(b)) on  $|Q_{mp_1}|^2$  at location  $(m/B, p_1/A_1)$  of the temporal-spatial frequency domain. For the lexicographic  $TS(1)$  sequence and the bit-reversed  $TS(1)$  sequence with  $M_1 = 16$ , the  $|Q_{mp_1}|$  for the lexicographical sequence and the bit-reversed sequence are shown in Figure 3.9. The  $|Q_{mp_1}|$  weights for the lexicographical sequence is large for term close to  $m = 0, p_1 = 0$ . This results in frequency translates that severely aliased with the  $m = 0, p_1 = 0$  frequency content term.

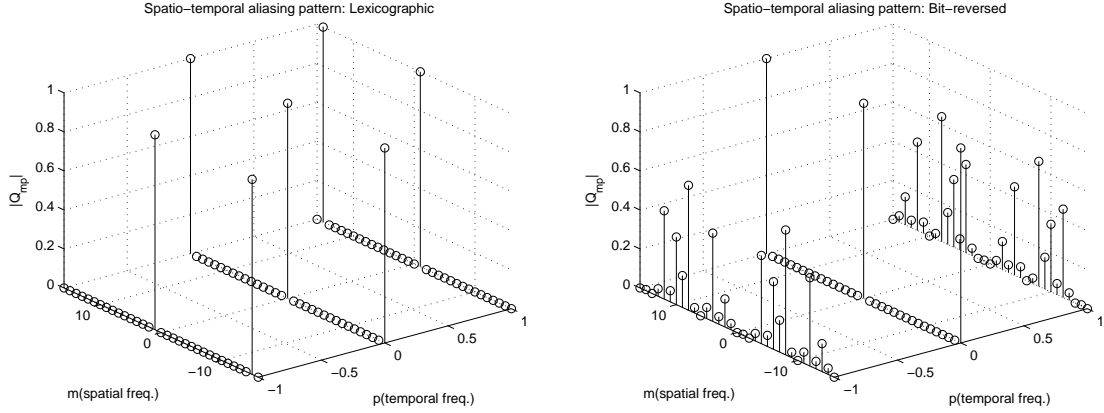


Figure 3.9: Spectral replication pattern  $|Q_{mp_1}|$  for lexicographical (left) and bit-reversed sequences (right).

Conversely, the  $|Q_{mp_1}|$  weights are more uniformly distributed for the bit-reversed case, where the small weights close to the  $m = 0, p_1 = 0$  frequency term result in negligible aliasing. The consequence of this result is that for signals with high temporal-tonal bandwidth, the lexicographical  $TS(1)$  sampling strategy will have more severe aliasing compared to the bit-reversed  $TS(1)$  sampling strategy. Hence the reconstruction from the bit-reversed  $TS(1)$  sampling sequence will be more tolerant to high temporal-tonal bandwidth signal. This means that the effectiveness of the  $TS(1)$  sampling depends heavily on the sequence used (i.e. bit-reversed or lexicographic) to sample the temporal-spatial signal.

The lexicographic and bit-reversed  $TS(1)$  sampling mentioned previously are somewhat arbitrary. Indeed there are several key problems in using these time-sequential sampling sequences, given as follows:

1. The lexicographic and bit-reversed time-sequential sampling sequences are limited to sampling of one spatial sample (i.e.  $n = 1$ ) at each sampling instance,  $k$ . A general  $TS(n)$  sampling for any  $n \geq 1$  would be desirable to take advantage of the availability of multiple  $n$  sensors.
2. Both the lexicographic and bit-reversed time-sequential sampling sequences will not completely avoid aliasing of the compact spectral support for sampling beyond the temporal Nyquist rate. A non-aliasing time-sequential sampling approach will be desirable to improve the sensing performance, when the time-sequentially sampled signal is eventually reconstructed.
3. With the lexicographic and bit-reversed time-sequential sampling, there is no approach

to determine both the time resolution (given by the inter-sampling time,  $T$ ) and spatial resolution,  $\mathbf{A} \cdot / \mathbf{M}$  (given by the number of discretized spatial points,  $\mathbf{M}$ ) so that aliasing of the assumed compact spectral support is avoided.

These deficiencies motivate the need for designing the  $TS(n)$  sampling approach that; 1) avoids aliasing of the compact spectral support in the sampled signal temporal-spatial frequency domain and; 2) for all feasible non-aliasing  $TS(n)$  sampling, we would like to find one that has the largest inter-sampling time  $T$ . In general this can be achieved by rearranging the replicas of the compact spectral support of  $\Omega(f, u_1)$ ,  $\Theta(W, U_1)$  in the sampled signal's frequency domain such that no aliasing occurs (i.e no overlapping of the replicas of  $\Theta(W, U_1)$ ). Then, the appropriate time-sequential sampling sequence, time resolution and spatial resolution can be derived from this new non-aliasing arrangement. The actual process of achieving this basic idea is realized in the lattice theoretic framework with an optimization algorithm for any  $1 + N$  dimensional temporal-spatial signal. This is given in Chapter 4 and Chapter 5.

### 3.3 Reconstruction of Time-Sequentially Sampled Signal

Assuming that no aliasing occurs, time-sequentially sampled signals are traditionally reconstructed by low-pass temporal-spatial filtering based on the assumed signal spectral support. This approach, however, is problematic in feedback control application for the following reasons:

1. The zero-phase ideal low pass filter cannot be implemented causally.
2. When the reconstructed temporal-spatial signal is used for feedback control, the frequency content of the temporal-spatial signal will be significantly increased due to the control input. This induces undue aliasing, while the inherent process dynamics actually remains the same. However, it is not apparent how to incorporate the known information about the control input in the low-pass reconstruction technique.

These difficulties can be resolved using an alternate approach based on Kalman filtering.

#### 3.3.1 Reconstruction via a Kalman filter

From (3.7) let  $\Psi[\bar{\mathbf{v}}(k)] := \hat{\phi}(\mathbf{I} + \Delta(k)\mathbf{W}_v)\bar{\mathbf{v}}(k)$ . Substituting (3.11) into equation (3.7) we have:

$$\bar{\mathbf{e}}(k) = \Psi[\bar{\mathbf{v}}(k)] + \mathbf{F}\mathbf{x}_d(k) \quad (3.32)$$

where  $\mathbf{F} = [\mathbf{F}_\ell, \mathbf{F}_h] \in \mathbb{R}^{NM_t \times 2NM_t}$  and  $\mathbf{x}_d(k) = [\mathbf{x}_\ell(k), \mathbf{x}_h(k)]^T \in \mathbb{R}^{2NM_t}$ . Augmenting (3.12) and (3.13) we have:

$$\begin{aligned}\mathbf{x}_w(k+1) &= \mathbf{A}_w \mathbf{x}_w(k) + \mathbf{B}_w \mathbf{w}(k) \\ \mathbf{x}_d(k) &= \mathbf{C}_w \mathbf{x}_w(k) + \mathbf{D}_w \mathbf{w}(k)\end{aligned}\quad (3.33)$$

where  $\mathbf{x}_w(k) = [\mathbf{x}_1(k), \mathbf{x}_2(k)]^T \in \mathbb{R}^{NM_t(n_{f_1} + n_{f_2})}$ ,  $\mathbf{w}(k) = [\mathbf{w}_1(k), \mathbf{w}_2(k)]^T \in \mathbb{R}^{2NM_t}$  and

$$\begin{aligned}\mathbf{A}_w &= \begin{pmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{pmatrix} \in \mathbb{R}^{NM_t(n_{f_1} + n_{f_2}) \times NM_t(n_{f_1} + n_{f_2})} \\ \mathbf{B}_w &= \begin{pmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{pmatrix} \in \mathbb{R}^{NM_t(n_{f_1} + n_{f_2}) \times 2NM_t} \\ \mathbf{C}_w &= \begin{pmatrix} \mathbf{C}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 \end{pmatrix} \in \mathbb{R}^{2NM_t \times NM_t(n_{f_1} + n_{f_2})} \\ \mathbf{D}_w &= \begin{pmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{pmatrix} \in \mathbb{R}^{2NM_t \times 2NM_t}\end{aligned}$$

where  $\mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1, \mathbf{D}_1$  are given in (3.12) and  $\mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2, \mathbf{D}_2$  are given in (3.13).

From the definition given in (3.20), we can define the signal,

$$\begin{aligned}\mathbf{y}(k) &:= \mathbf{C}_\alpha(k) \bar{\mathbf{e}}(k) - \mathbf{C}_\alpha(k) \Psi[\bar{\mathbf{v}}(k)] + \mathbf{n}(k) \\ &= [\mathbf{C}_\alpha(k) \mathbf{F} \mathbf{C}_w] \cdot \mathbf{x}_w(k) + [\mathbf{C}_\alpha(k) \mathbf{F} \mathbf{D}_w] \cdot \mathbf{w}(k) + \mathbf{n}(k)\end{aligned}\quad (3.34)$$

where both  $\mathbf{w}(k)$  and  $\mathbf{n}(k)$  are zero-mean white noise sequences with covariance  $\mathbf{R}_{ww} = \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{pmatrix} \in \mathbb{R}^{2NM_t \times 2NM_t}$  and  $\mathbf{R}_{nn} \in \mathbb{R}^{nN \times nN}$  respectively.  $\mathbf{y}(k) \in \mathbb{R}^{nN}$  in (3.34) can be treated as the measurement for the observer.

Notice that (3.33) and (3.34) is a linear time varying system because of  $\mathbf{C}_\alpha(k)$  in (3.34). If  $\alpha(k) = \alpha(k + M_p)$ , then (3.32) to (3.34) is a  $M_p$  periodic system. Under very general conditions, the  $M_p$  periodic linear system is observable and therefore admits a  $M_p$ -periodic Kalman filter to form the estimate  $\hat{\mathbf{d}}(k)$  of the disturbance  $\bar{\mathbf{d}}(k)$  [39]:

$$\begin{aligned}\hat{\mathbf{x}}_w(k+1) &= \mathbf{A}_c(k) \hat{\mathbf{x}}_w(k) + \mathbf{B}_c(k) \mathbf{y}(k) \\ \hat{\mathbf{d}}(k) &= \tilde{\mathbf{F}} \cdot \hat{\mathbf{x}}_w(k)\end{aligned}\quad (3.35)$$

where

$$\begin{aligned}\mathbf{A}_c(k) &= \mathbf{A}_w(\mathbf{I} - \mathbf{L}(k)\tilde{\mathbf{C}}_\alpha(k)\tilde{\mathbf{F}}) \\ \mathbf{B}_c(k) &= \mathbf{A}_w\mathbf{L}(k) \\ \tilde{\mathbf{F}} &= \mathbf{F}\mathbf{C}_w\end{aligned}$$

and  $\mathbf{L}(k)$  is the periodic Kalman filter gain obtained by solving the periodic Riccati equation:

$$\begin{aligned}\bar{\mathbf{P}}(k+1) &= \mathbf{A}_w[\bar{\mathbf{P}}(k) - \mathbf{L}(k)\tilde{\mathbf{C}}_\alpha(k)\tilde{\mathbf{F}}\bar{\mathbf{P}}(k)]\mathbf{A}_w^T + \mathbf{B}_w\mathbf{R}_{ww}\mathbf{B}_w^T \\ \bar{\mathbf{P}}(k) &= \bar{\mathbf{P}}(k + M_t)\end{aligned}\tag{3.36}$$

$$\mathbf{L}(k) = \bar{\mathbf{P}}(k)\tilde{\mathbf{F}}^T\tilde{\mathbf{C}}_\alpha^T(k) \left[ \mathbf{R}_{nn} + \tilde{\mathbf{C}}_\alpha(k)\tilde{\mathbf{F}}\bar{\mathbf{P}}(k)\tilde{\mathbf{F}}^T\tilde{\mathbf{C}}_\alpha^T(k) \right]^{-1}\tag{3.37}$$

The solution to the periodic Riccati equation can be obtained by iterating (3.36) until  $\bar{\mathbf{P}}(k)$  converges, or equivalently by solving for the steady state solution using approaches, such as lifting, that can be found in [40, 41, 42]. The estimate of the temporal-spatial error  $\hat{\mathbf{e}}(k)$  can be reconstructed as:

$$\hat{\mathbf{e}}(k) = \Psi[\mathbf{v}(k)] + \hat{\mathbf{d}}(k)\tag{3.38}$$

### 3.3.2 Analysis of time-sequential sampling and reconstruction via a Kalman filter

In this section, the  $M_p$ -periodic Kalman reconstruction filter (3.35) with respect to its input spectrum will be analyzed. While the frequency response of a periodic system is not as straightforward as for a time invariant system, there are several attempts in the literature, such as harmonic transfer function [43, 44, 45] and lifting [46]. In this dissertation, the fundamental Floquet theory for periodic systems [47] was applied to decompose the filter's input/output relationship to gain insights into how the spectrum is transformed.

According to Floquet theory, the transition matrix  $\Phi(k, k_0) \in \mathbb{R}^{NM_t(n_{f_1}+n_{f_2}) \times NM_t(n_{f_1}+n_{f_2})}$  for the  $M_p$ -periodic discrete time system (3.35) can be written as:

$$\Phi(k, k_0) = \mathbf{P}(k)\mathbf{\Lambda}^{k-k_0}\mathbf{P}^{-1}(k_0)$$

where  $\mathbf{\Lambda} := \Phi(M_p, 0)^{1/M_p}$  is a constant matrix and  $\mathbf{P}(k) \in \mathbb{R}^{NM_t(n_{f_1}+n_{f_2}) \times NM_t(n_{f_1}+n_{f_2})}$  is a  $M_p$ -periodic matrix sequence given by:

$$\mathbf{P}(k) = \begin{cases} \Phi(k, 0)\mathbf{\Lambda}^{-k} & \text{for } k \in [0, M_p - 1] \\ \mathbf{P}(k - M_p) & \text{when } k \geq M_p \end{cases}.$$



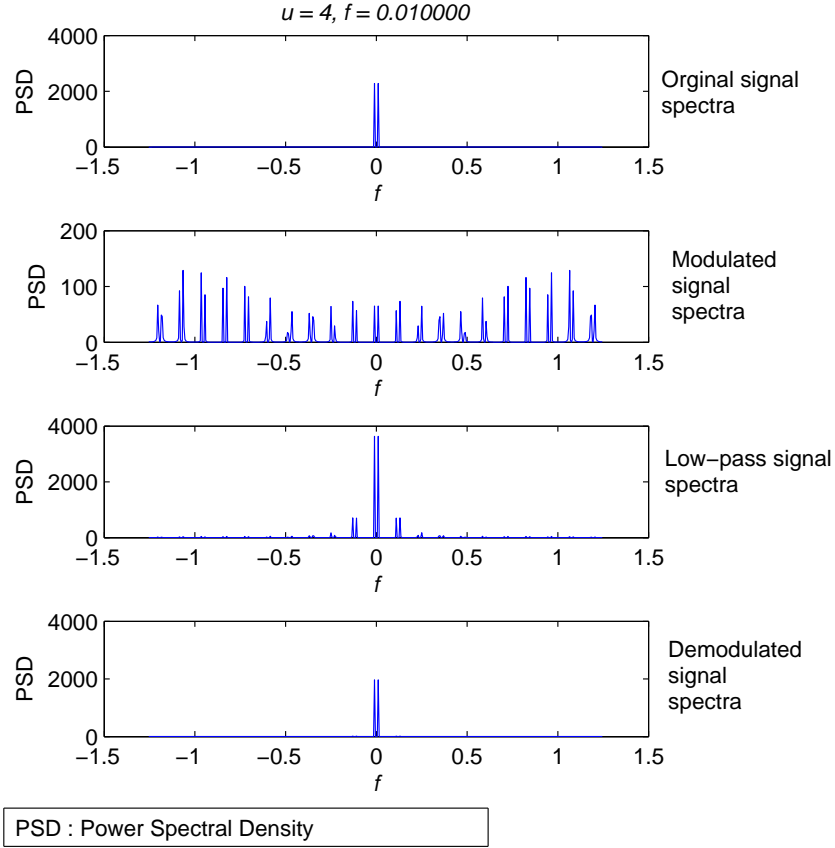


Figure 3.10: Transformation of a sinusoidal  $1+N = 2$  temporal-spatial ( $u_1 = 4$  [cycles/ton],  $f = 0.01$  [Hz]) disturbance input through the three processes in the Kalman filter. From top to bottom are the temporal spectra for a particular tone: 1) original disturbance signal; 2) transformed by  $\tilde{\mathbf{F}}\mathbf{P}(k+1)^{-1}\mathbf{B}_c(k)$ ; 3) after filtering by  $\mathbf{\Lambda}$ ; 4) demodulation by  $\tilde{\mathbf{F}}\mathbf{P}(k)\tilde{\mathbf{F}}^{-1}$ .

Applying this decomposition to the convolution formula for (3.35), we have:

$$\hat{\mathbf{d}}(k) = [\tilde{\mathbf{F}}\mathbf{P}(k)\tilde{\mathbf{F}}^{-1}] \cdot \left\{ \tilde{\mathbf{F}} \sum_{k'=k_0}^{k-1} \mathbf{\Lambda}^{k-k'-1} \tilde{\mathbf{F}}^{-1} [\tilde{\mathbf{F}}\mathbf{P}(k'+1)^{-1}\mathbf{B}_c(k') \cdot \mathbf{y}(k')] \right\} \quad (3.39)$$

Equation (3.39) shows that, in the original disturbance coordinates, the Kalman filter operates on the  $\mathbf{y}(k)$  in three steps:

1. multiplication by the periodic static operator -  $\mathbf{y}(k) \rightarrow \tilde{\mathbf{F}}\mathbf{P}(k+1)^{-1}\mathbf{B}_c(k) \cdot \mathbf{y}(k)$ .
2. time invariant linear filtering by the filter -  $\mathbf{y}(k) \rightarrow \tilde{\mathbf{F}}\mathbf{x}(k)$  with  $\mathbf{x}(k+1) = \mathbf{\Lambda}\mathbf{x}(k) +$

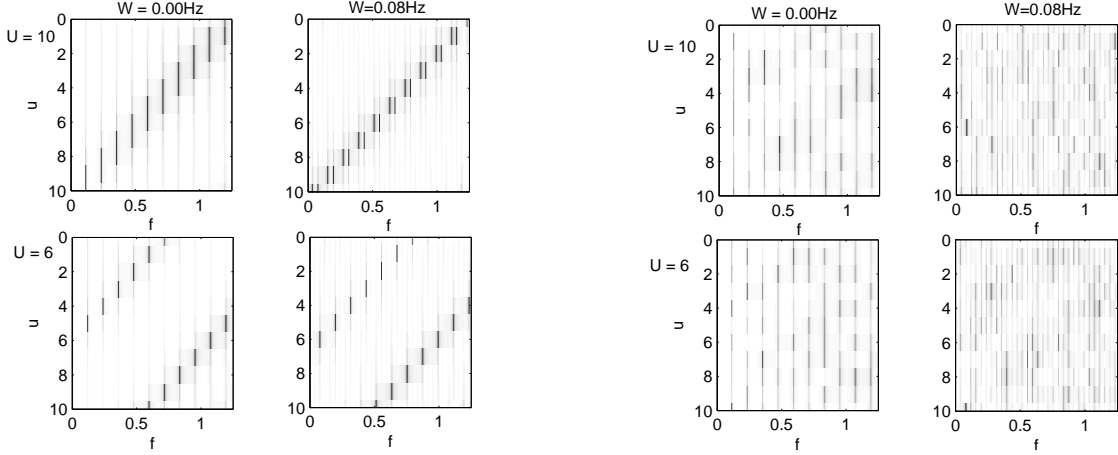


Figure 3.11: Temporal-tonal power-spectral density for various single tone input frequencies ( $u_1$  [cycles/tone],  $f$  [Hz]) after the first transformation of the reconstruction filter for the **lexicographic**(left) and **bit-reversed** (right)  $TS(1)$  sampling.

$$\tilde{\mathbf{F}}^{-1}\mathbf{y}(k)$$

3. multiplication by another periodic static operator -  $\mathbf{y}(k) \rightarrow \tilde{\mathbf{F}}\mathbf{P}(k)\tilde{\mathbf{F}}^{-1}\mathbf{y}(k)$ .

The three steps are illustrated for a low single temporal-tonal frequency disturbance input signal (at one tone  $u_1 = 4$  [cycles/tone],  $f = 0.01$  [Hz]) in Figure 3.10. Step 1 modulates the input signal by duplicating the weighted copies of the original signal according to the spectrum of  $\tilde{\mathbf{F}}\mathbf{P}(k'+1)^{-1}\mathbf{B}_c(k')$ . Step 2 represents a low pass filtering according to the matrix  $\mathbf{A}$  which eliminates nearly all but the original signal frequency and one extra copy. Step 3 corresponds to a demodulation due to the multiplication by  $\tilde{\mathbf{F}}\mathbf{P}(k)\tilde{\mathbf{F}}^{-1}$ . Interestingly, the demodulation folds the high frequency copy of the signal after step 2 back to the low frequency, resulting in a reconstruction with minimal error.

The modulation effect of step 1 can also be thought of as modulating the spectral content of the map  $\tilde{\mathbf{F}}\mathbf{P}(k+1)^{-1}\mathbf{B}_c(k)$  by the input signal. The pattern in which the single tonal-temporal frequency of the disturbance signal is modulated by this map is determined by the spectral replication pattern ( $Q_{mp1}$ ) of the sampling pattern shown in Figure 3.9. Figure 3.11 shows the spectral contents of the signal after the first transformation for different single tonal-temporal input frequencies for the lexicographic and bit-reversed time-sequential sampling sequences. As the temporal frequency increases (left to right), copies of the spectral content of the map approach each other on the temporal frequency domain (horizontally); and similarly as the tonal frequency increases (bottom to top), copies of the spectrum of the map approach each other on the tonal frequency domain (vertical). Notice that for

the lexicographical sampling sequence, the frequency replications are very distinct and follow a structured pattern; whereas for the bit-reversed sampling sequence, the frequency replications are much more diffused. Therefore for the class of signal with a compact ellipsoidal tonal-temporal spectral support, the modulation process (step 1) with spectral support larger than the corresponding Nyquist frequencies would result in a very distinct and intense overlap with the lexicographic sampling sequence; whereas for the bit-reversed sequence, the overlap process is much more diffused and gentle.

### 3.4 Case Study

In this case study, the TRC is considered i.e. a  $1+N = 2$  dimensional temporal-spatial signal is considered. The TRC is assumed to be well represented by its values at  $M_1 = M = 21$  tones so that  $\bar{\omega}(k) \in \mathbb{R}^{M=21}$  (see equation (3.2)). For simplicity, we assume that in (3.7), the effect of controls have been subtracted out i.e.  $\bar{\mathbf{v}}(k) = \mathbf{0}$  and consider behavior of system without plant perturbation i.e.  $\Delta(k) = 0$ . Furthermore, we use  $T = 0.4$  [s],  $A = 1$ ,  $1/MT = 0.12$  [Hz] so that the tonal-temporal Nyquist frequencies are  $(u_N, f_N) = (10.5$  [cycles/tone],  $0.06$  [Hz]).

Under very general conditions the  $M$ -periodic linear system is observable and therefore admits a  $M$ -periodic Kalman filter to form the estimate of the TRC at time instance,  $k$  given by  $\hat{\mathbf{e}}(k) + \bar{\omega}^*$  where  $\bar{\omega}^*$  is the nominal TRC. To test the Kalman filter we generate the  $\bar{\mathbf{e}}(k) = \bar{\mathbf{d}}(k)$  (since we assume  $\bar{\mathbf{v}}(k) = \mathbf{0}$ ) to have a compact spectral support  $\{(u_1, f) | (u_1/U_1)^2 + (f/W)^2 \leq 1\}$  where  $U_1$  and  $W$  are the highest spatial and temporal frequencies in  $\bar{\omega}(k) = \bar{\mathbf{e}}(k) + \bar{\omega}^*$  using (3.11). The ratio of the mean squared reconstruction error to the mean squared deviation from the nominal TRC,

$$\Phi := \frac{E[\|\bar{\mathbf{e}}(k) - \hat{\mathbf{e}}(k)\|_2^2]}{E[\|\bar{\mathbf{e}}(k)\|_2^2]}$$

for each pair of support size  $(U_1, W)$  can be used to compare the filter performance at different temporal-spatial support frequencies. Figure 3.12 shows the performance ratio,  $\Phi$  of the Kalman filters using lexicographical, and bit-reversed sampling sequences. Notice that at low tonal-temporal frequencies, both filters reconstruct well. Noticeably, at high tonal frequencies (above the tonal-temporal Nyquist frequency), the Kalman filter for the lexicographical sequence reconstructs worse than the one for the bit-reversed sequence. By comparing the two performance error map, Figure 3.13 shows the area in the tonal-temporal domain where we can get better reconstruction performance with the bit-reversed sampling pattern. The maximum difference between the lexicographic and bit-reversed sampling sequences occur at  $(U_1, W) = (3.0, 0.31)$ . This result can also be seen from the actual

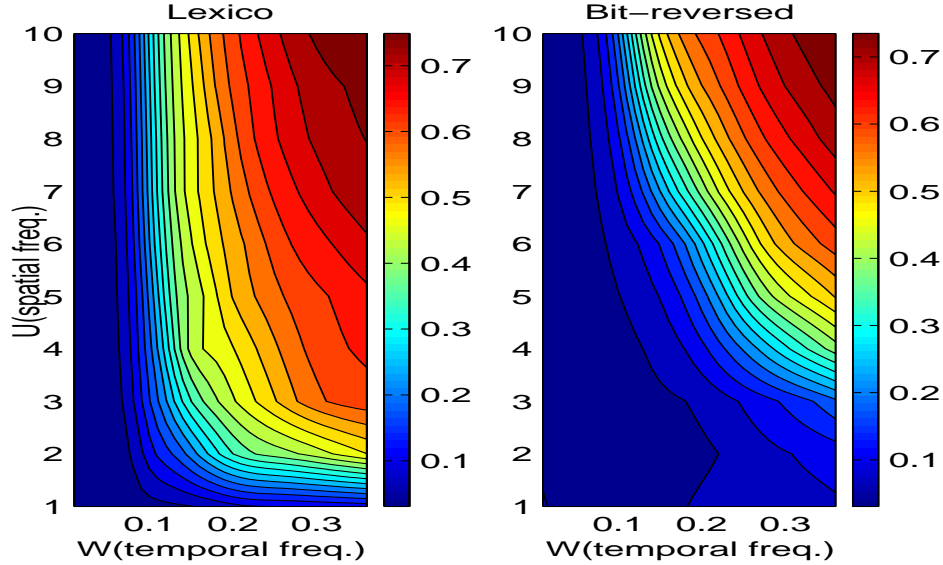


Figure 3.12: Ratio of mean squared reconstruction errors to mean squared deviation from the nominal TRC,  $\Phi$  using lexicographical and bit-reversed sequences at various tonal-temporal support frequencies  $(U_1, W)$ .

reconstruction error in Figure 3.14. The clear difference between the performances of the two sampling strategies can be explained by the differences in the severity of aliasing as defined by the aliasing weights in Figure 3.9.

The mean-squared error does not convey the worst case error however. The worst case errors (after convergence) for the two sampling strategies are shown in Figs. 3.15-3.16. They show that although perfect reconstruction is achieved at low temporal frequencies, the performance degrades as the tonal and temporal support frequencies  $(U_1, W)$  increase. It also shows a better TRC reconstruction using the bit-reversed sequence especially at  $U_1 = 6.0$ .

### 3.5 Summary

Time-sequential sampling (i.e.  $TS(1)$  sampling using lexicographic and bit-reversed sequences) is proposed in this dissertation to increase the printer's ability to sense the high dimensional time varying tone reproduction curve (TRC) or color reproduction characteristics (CRC) function, which is useful for control of color consistency. A Kalman filter based method has been developed and shown to be effective to reconstruct the original function from the time-sequential samples. Using Floquet theory, it is shown that the periodic Kalman reconstruction filter consists of a cascade of modulation, low-pass filtering, and de-

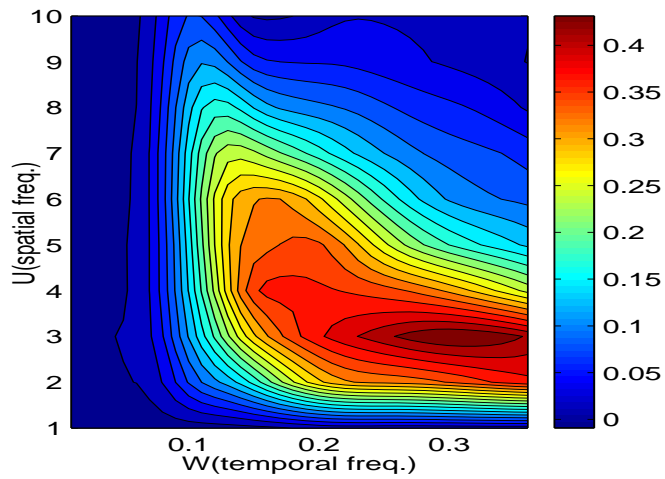


Figure 3.13: Differences of  $\Phi$  between bit-reversed and lexicographic sequences at various tonal-temporal support frequencies ( $U_1, W$ ).

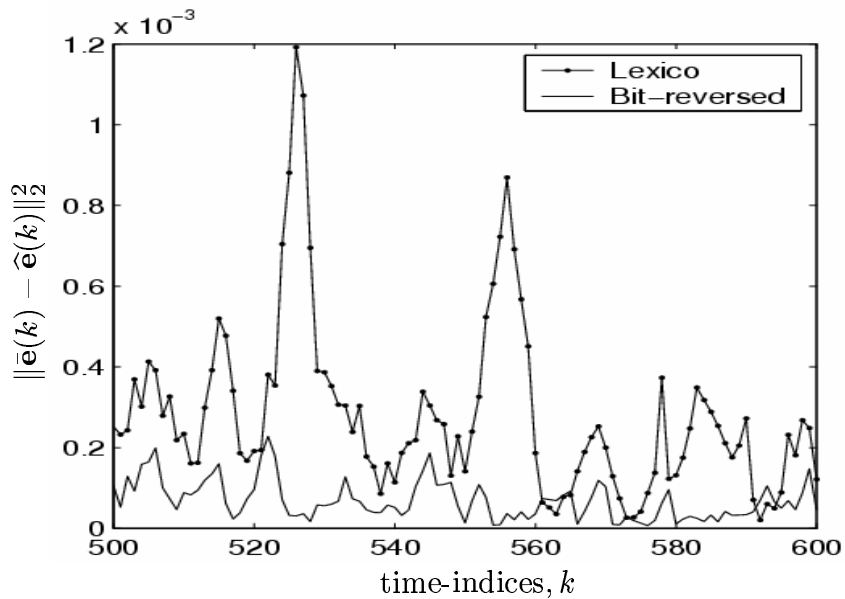


Figure 3.14: Reconstruction error ( $\|\bar{\mathbf{e}}_1(k) - \hat{\mathbf{e}}_1(k)\|_2^2$ ) as a function of time using lexicographical and bit-reversed sequences at  $(U_1, W) = (3.0, 0.31)$

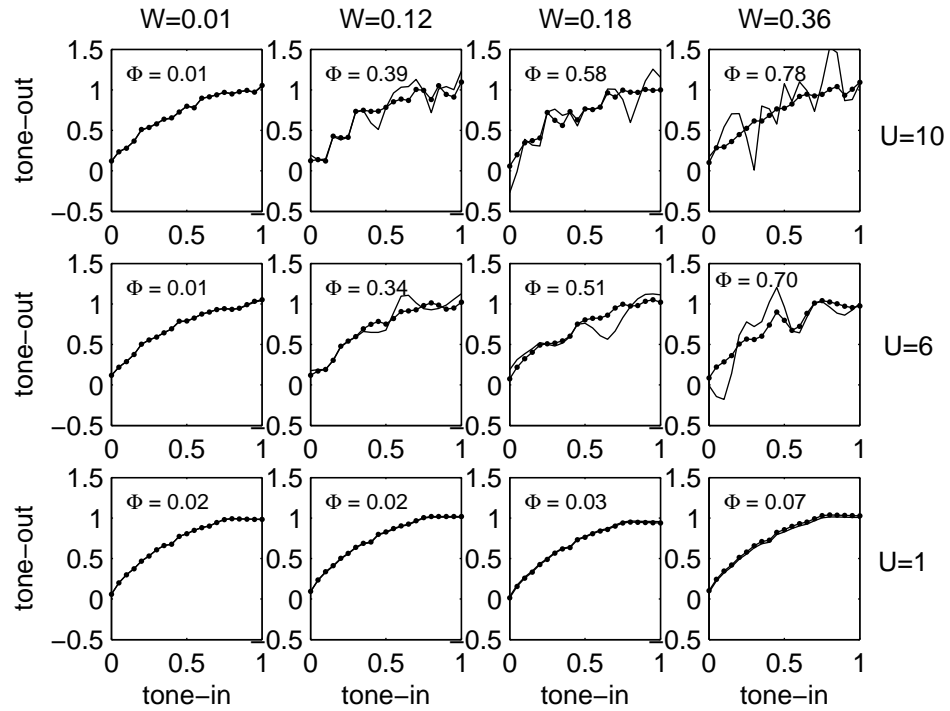


Figure 3.15: Input TRC (consisting of a typical TRC with various tonal-temporal support frequencies ( $U_1, W$ ) imposed on it) and its worst case steady state reconstruction, for Kalman filter designed for the lexicographic strategy (dotted line:reconstructed signal, line:input signal).

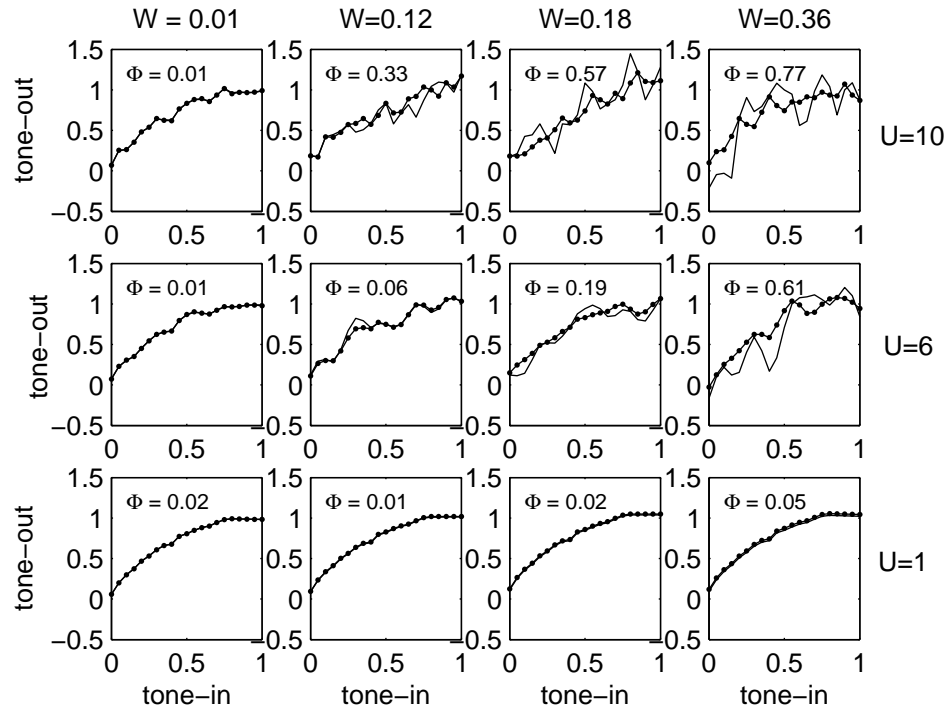


Figure 3.16: Input TRC (consisting of a typical TRC with various tonal-temporal support frequencies ( $U_1, W$ ) imposed on it) and its worst case steady state reconstruction, for Kalman filter designed for the bit-reversed strategy (dotted line:reconstructed signal, line:input signal).

modulation transformation. It shows that the performance of the time-sequential sampling scheme depends both on the signal model and the sampling strategy. The dependence of sampling strategy and performance will be discussed further in the next chapter, to assist in the design of the  $TS(n)$  sampling.



## Chapter 4

# Lattice Theoretic Framework for Time-Sequential Sampling

In the previous chapter, the lexicographic and bit-reversed  $TS(1)$  sampling sequence were presented. However, both of these sampling sequences do not completely avoid aliasing and neither the spatial (tonal) resolution nor the inter-sampling time was optimized. In addition, only one spatial sampling point is considered at each time step,  $k$ . In this chapter, we address these deficiencies. Specifically we find the  $TS(n)$  sampling that fulfills the following conditions:

- (c1) Requires only  $n$  small number of spatial samples at each time-step,  $k$ .
- (c2) The desired  $TS(n)$  sampling does not result in aliasing of the assumed compact spectral support,  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  (see (3.8) and (3.10)). This ensures minimal reconstruction error.
- (c3) For known compact spectral support of the temporal-spatial signal,  $\Theta(W, \mathbf{U})$  and maximum spatial resolution,  $A_i/(M_{max})_i$  at each  $i$ -the spatial dimension (see Figure 3.3), the desired  $TS(n)$  sampling yields the maximum inter-sampling time  $T$  without violating condition (c2).

Conditions (c1), (c2) and (c3) represent the objectives for sensing the TRC and CRC for their stabilization. Fulfilling these conditions are critical here because the high cost and dimensionality of sensing dictates that sensing the TRC and CRC be done infrequently. Achieving these sensing objectives require the determination of:

1. Which  $n$  samples to take and in what order i.e. what is  $\boldsymbol{\alpha}(k)$  defined by (3.17)-(3.18)?
2. How to choose the right inter-sampling time,  $T \in \mathbb{R}^+$ ?
3. Let  $(M_{max})_i$  gives the maximum spatial discretization points at each  $i$ -th spatial dimension. If each  $i$ -th spatial dimension is discretized into  $M_i$  points, then what should  $M_i \leq (M_{max})_i$  be?

These parameters and the known  $N$ -dimensional spatial range given by vector  $\mathbf{A} \in \mathbb{R}^N$  give the quadruple  $(\boldsymbol{\alpha}(k), T, \mathbf{A}, \mathbf{M})$  that defines the desired  $TS(n)$  sampling. Determination of these parameters i.e. the desired  $TS(n)$  sampling through a design process are needed for effective and efficient sensing of the time-varying TRC/CRC or in general, arbitrary temporal-spatial signal. On top of being used to sense the high dimensional time-varying TRC/CRC, the  $TS(n)$  sampling design may also be beneficial for other spatially distributed systems such as paper making, plastic film extrusion, steel rolling, coating and laminating. In these processes, the film properties with spatial dependencies are measured via an array of sensors or via a scanning sensor that moves back and forth in the cross direction. Obviously without sensing approach like the  $TS(n)$  sampling, sensing the multiple dimension simultaneously is very costly. While having a well-designed  $TS(n)$  sampling is well founded, previous works on time-sequential sampling (such as the lexicographic and bit-reversed time-sequential sampling sequence discussed in the previous chapter) do not sufficiently provide a means to achieve the desired sensing requirements as listed by conditions (c1), (c2) and (c3).

Allebach [11, 38], the first to consider the time-sequential sampling problem, shows that the severity of the aliasing of the time-sequentially sampled signal is sequence dependent. Hence the amount of aliasing can be controlled by changing the sampling sequence. This motivates Allebach to study the design of the sampling sequence [12] which minimizes the mean-squared reconstruction error for the case of white ellipsoidally bandlimited stochastic signals with rectangular spatial support. This approach suffers from several problems:

- (i) This design requires the designer to pick the spatial locations for the sample points *à priori*. In other word, the spatial discretization determined by  $M_i \in \mathbb{Z}^+$  needs to be predetermined. Thus they eliminate important degree of freedom from the design process. Specifically, the available sensing resolution cannot be imposed in the design process to optimize the inter-sampling time,  $T \in \mathbb{R}^+$ .
- (ii) The design considers only the case of two-spatial dimension i.e.  $N = 2$ .

- (iii) To avoid an infeasible combinatorial search over all possible sequences, this solution was suboptimal, involving an *ad-hoc* search over a possible subset of possible sequences, which were generated using the Chinese Remainder Theorem.
- (iv) Finally, consideration is only given for one spatial samples at each time-step i.e.  $n = 1$ .

In the interest of creating a better design technique, Willis and Bresler [13, 14] proposed an optimal design solution by restricting the  $TS(1)$  sampling sequence to lie on a lattice in the temporal-spatial domain. Their proposed optimal  $TS(1)$  sampling design framework eliminates the need to pre-specify the spatial locations for sampling and is extensible to spatial signals of any dimension and spatial support. The restriction to lattice patterns was motivated by the ease and simplicity with which lattice theory was applied to the general multidimensional sampling problem [17]. They gave a rather complete theoretical analysis of the  $TS(1)$  sampling in the lattice theoretic framework for any arbitrary temporal-spatial frequency support. In [16], Willis and Bresler demonstrated the actual design process to obtain the optimal  $TS(1)$  sampling for the sampling of time-varying tomographic images. This work looked into the cross-sectional tomographic imaging that varies with time (e.g. time-varying cross-sectional cardiac imaging) i.e.  $1 + N = 3$  dimensional temporal-spatial signal. The nature of tomographic scan imaging, makes it possible to reduce this to a problem of designing the  $TS(1)$  sampling for  $1 + N = 2$  dimensional signal. However, there are several difficulties:

- (i) The proposed design approach in [16], is demonstrated only for the simple case of designing the  $TS(1)$  sampling for  $1 + N = 2$  dimensional signal. While huge computational complexity is avoided, extending the proposed design approach to higher dimensional signal is difficult because the heuristic design method requires manual “by inspection” check for  $TS(1)$  sampling and frequency content packing. Additionally, the heuristic design method may not yield the best possible  $TS(1)$  sampling design, given the available sensing resources and capabilities. An example of this proposed design approach given by [16] is demonstrated in Appendix C.1.
- (ii) Condition given in [13, 14] (also stated by Theorem 4.3 in our time-sequential sampling framework) can be easily used to check if a known lattice is a non-aliasing  $TS(1)$  sampling lattice. However, it is difficult to enforce these conditions on an unknown  $TS(1)$  sampling lattice to fulfill the requirements of conditions (c1), (c2) and (c3).
- (iii) Similar to Allebach’s approach, consideration is only given for one spatial samples at each time-step i.e.  $n = 1$ .

In this dissertation, we first proposed an optimization approach for finding the optimal  $TS(1)$  sampling that fulfills the conditions (c1), (c2) and (c3). The optimization process eliminates the need for manual “by inspection” check. We then construct the  $TS(n)$ ,  $n > 1$  sampling using the optimal  $TS(1)$  sampling with an appropriately scaled spatial domain. Both of these make up important *contributions* of this dissertation on yielding a self-contained algorithm to design effective  $TS(n)$  sampling for sampling temporal-spatial signals of any  $1 + N$  dimension. Such algorithm not only enables ease of use, putting the guess works out of the design process but also offers an opportunity to adaptively update the  $TS(n)$  sampling for changing frequency contents of the temporal-spatial signal.

Since the search space for the inter-sampling time,  $T$ ; the number of discretization points,  $M_i$  at each  $i$ -th spatial dimension and the  $TS(1)$  sampling sequences could potentially be very large, finding the optimal solution is exceedingly complex and computationally prohibitive. To make the problem tractable, the  $TS(1)$  sampling sequences are restricted to be a  $1 + N$  dimensional lattice. This restriction is not severe. Although not all  $TS(1)$  sampling are lattices, we can show that in our sampling framework as given in Chapter 3, an optimal  $TS(1)$  lattice will be as good as an unconstrained optimal  $TS(1)$  sampling provided that the temporal-spatial signal is not undersampled temporally by more than a factor of 2.

In addition, while not all  $TS(1)$  sampling are lattices, not all lattices form  $TS(1)$  sampling. The Hermite Normal Form is used to define unique lattices as proposed by [13, 14]. The lower-triangular Hermite Normal Form gives useful geometric insights on the lattice that can be used to define the  $TS(1)$  sampling lattice. A sufficient and necessary condition for a lattice to be a  $TS(1)$  sampling lattice is given in [13]. The lattice theoretic framework is useful because it allows us to convert the temporal-spatial domain requirements into the frequency domain and vice versa. A sufficient condition can then be defined to characterize all non-aliasing lattices. In this dissertation, we parameterized the space of lattices by the inter-sampling time  $T$ , the number of spatial discretization at each dimension  $M_i \in \mathbb{Z}^+$  and the spatial range  $A_i \in \mathbb{R}^+$ . The optimization problem for finding the  $TS(1)$  sampling lattice that avoids aliasing and maximizes the inter-sampling time can then be carried out with respect to these parameterizations. It will be shown that, by imposing the  $TS(1)$  sampling condition and the non-aliasing condition on this parameterizations, the space of all feasible non-aliasing  $TS(1)$  sampling lattices is compact (i.e. finite cardinality). Hence optimization procedures that guarantee to terminate in finite time can be proposed.

Although finite, the cardinality of the set of feasible non-aliasing  $TS(1)$  sampling lattices can be very large. This results in optimization process with high computational complexity. For signal with high dimension this is particularly problematic. Therefore it is important to

restrict the set of feasible non-aliasing  $TS(1)$  sampling lattices for optimization. Approaches for reducing the computational complexity by restricting the search space are proposed.

The proposed  $TS(n)$  sampling fulfills criteria (c1) and gives an inter-sampling time,  $T$  that linearly scales the inter-sampling time of the optimal  $TS(1)$  sampling design with the  $n$  number of samples. While this sub-optimal  $TS(n)$  sampling lattice design does not necessarily fulfill criteria (c3), it allows us to arrive at a compromise between sensors hardware and inter-sampling time,  $T$  without loss in performance. In our application, this is of significant practical importance.

In section 4.1 the key properties of the lattice theoretic framework used for defining the time-sequential sampling lattices are given. Section 4.2 described the  $TS(n)$  sampling lattice. Section 4.3 described the parameterizations of the space of  $TS(1)$  sampling lattices by the inter-sampling time  $T \in \mathbb{R}^+$ , the number of spatial discretization at each dimension  $M_i \in \mathbb{Z}^+$  and the spatial range  $A_i \in \mathbb{R}^+$ . Since not all lattices form time-sequential sampling sequences, we obtain condition on the lattices that do form  $TS(1)$  sampling sequences. Furthermore, although not all  $TS(1)$  sampling are lattices, we can show that an optimal  $TS(1)$  lattice will be as good as an unconstrained optimal  $TS(1)$  sampling provided that the temporal-spatial signal is not undersampled temporally by more than a factor of 2 as presented in section 4.4. Building on this theoretical results, the optimization algorithm to realize the optimal  $TS(1)$  sampling and the  $TS(n)$  sampling will be presented in the next chapter.

#### 4.1 Lattice Theoretic Framework

Let  $\mathbf{V} \in \mathbb{R}^{(1+N) \times (1+N)}$  be a non-singular matrix. The lattice generated by  $\mathbf{V}$ ,  $LAT(\mathbf{V}) \subset \mathbb{R}^{(1+N)}$ , is a set of points given by:

$$LAT(\mathbf{V}) \equiv \{\boldsymbol{\xi} \in \mathbb{R}^{(1+N)} : \boldsymbol{\xi} = \mathbf{V}\mathbf{z}, \mathbf{z} \in \mathbb{Z}^{(1+N)}\} \quad (4.1)$$

The columns of  $\mathbf{V}$ ,  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{(1+N)}\}$  are the basis of a full-rank point lattice,  $LAT(\mathbf{V})$ , as shown in Figure 4.1.

We define a *unit cell* of a lattice  $LAT(\mathbf{V})$  as a set  $\Upsilon \subset \mathbb{R}^{(1+N)}$  (not necessarily connected) such that  $\mathbb{R}^{(1+N)}$  is the disjoint union of copies of  $\Upsilon$  centered on each lattice point:

$$(\Upsilon + \boldsymbol{\xi}) \cap (\Upsilon + \mathbf{y}) = \emptyset \text{ for } \boldsymbol{\xi}, \mathbf{y} \in LAT(\mathbf{V}), \boldsymbol{\xi} \neq \mathbf{y}$$

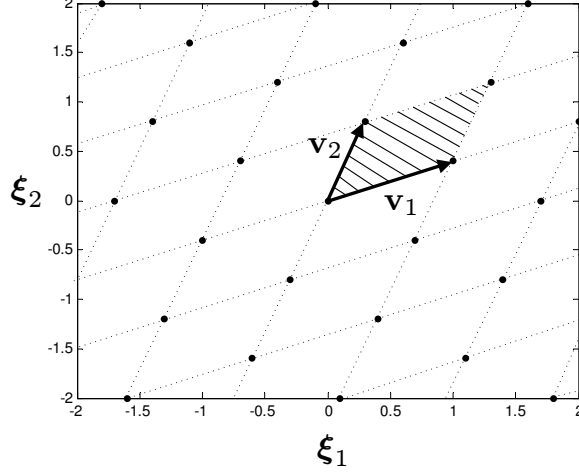


Figure 4.1: An example of two-dimensional point lattice  $LAT(\mathbf{V})[1 + N = 2]$  where  $\mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2]$ . The shaded area is the unit cell,  $\Upsilon$  with its volume given by  $\vartheta[LAT(\mathbf{V})]$

and,

$$\bigcup_{\boldsymbol{\xi} \in LAT(\mathbf{V})} (\Upsilon + \boldsymbol{\xi}) = \mathbb{R}^{(1+N)}$$

The hypervolume of a unit cell of a lattice  $LAT(\mathbf{V})$  is  $\vartheta[LAT(\mathbf{V})]$ . There are many possible choices for the unit cell of a lattice. One that is very convenient is the *fundamental parallelepiped* (see Figure 4.1) given by:

$$\Upsilon = \left\{ \sum_{i=1}^{1+N} \rho_i \mathbf{v}_i \mid 0 \leq \rho_i < 1 \text{ for } i = 1, \dots, 1 + N \right\}$$

Other key properties for lattices used throughout this chapter and next are listed as follow [13, 48, 49]:

**Property 4.1** *A multidimensional signal sampled on a lattice  $LAT(\mathbf{V})$  results in the replication of the original frequency content of the signal on a lattice  $LAT(\mathbf{V}^*)$  in the frequency domain, where*

$$\mathbf{V}^* = \mathbf{V}^{-T} \tag{4.2}$$

*Henceforth  $LAT(\mathbf{V}^*)$  is known as the polar lattice.*

This property is given by the following theorem:

**Theorem 4.1** *Let  $\Omega(f, \mathbf{u})$  be the power spectral density of the wide-sense stationary temporal-spatial signal,  $\boldsymbol{\omega}(t, \mathbf{x}) \in \mathbb{R} \times \mathbb{R}^N$ . The power spectral density of the sampled signal on a lattice*

at discrete locations,  $[t, \mathbf{x}]^T = \mathbf{V}\mathbf{z}$ ,  $\mathbf{z} \in \mathbb{Z}^{(1+N)}$  i.e. for the sampled signal

$$\omega_s(\mathbf{z}) = \omega(\mathbf{V}\mathbf{z})$$

the spectrum is given by:

$$\Omega_s(f, \mathbf{u}) = \frac{1}{\vartheta[LAT(\mathbf{V})]} \sum_{(m, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^N} \Omega \left( \begin{pmatrix} f \\ \mathbf{u} \end{pmatrix} - \mathbf{V}^{-T} \begin{pmatrix} m \\ \mathbf{p} \end{pmatrix} \right) \quad (4.3)$$

**Proof:** Let  $\boldsymbol{\xi} := [t, \mathbf{x}]^T$  and  $\boldsymbol{\nu} := [f, \mathbf{u}]^T$ . For wide sense stationary signals defined on a lattice  $LAT(\mathbf{V})$  we assume in the usual way that the process is a homogeneous random field with zero mean and auto-covariance function

$$R(\boldsymbol{\xi}) = E[\omega_s(\mathbf{y})\omega_s(\mathbf{y} + \boldsymbol{\xi})]$$

By Wiener-Khincine Theorem, the power spectral density of the random field is given by the Fourier transform of the auto-covariance function:

$$\Omega_s(\boldsymbol{\nu}) = \sum_{\boldsymbol{\xi} \in LAT(\mathbf{V})} R(\boldsymbol{\xi}) \exp^{-j2\pi\boldsymbol{\nu}^T \boldsymbol{\xi}} \quad (4.4)$$

$\Omega_s(\boldsymbol{\nu})$  is a periodic function with periodicity lattice equal to the polar lattice  $LAT(\mathbf{V}^*)$ , that is:

$$\Omega_s(\boldsymbol{\nu}) = \Omega_s(\boldsymbol{\nu} + \mathbf{r}), \forall \mathbf{r} \in LAT(\mathbf{V}^*)$$

This follows from the fact that  $\mathbf{r}^T \boldsymbol{\xi}$  is an integer for  $\boldsymbol{\xi} \in LAT(\mathbf{V})$  and  $\mathbf{r} \in LAT(\mathbf{V}^*)$ , by definition of the polar generator matrix given by (4.2). Because of this periodicity, the Fourier transform need only be specified for one unit cell,  $\Upsilon$  of the polar lattice. The inverse Fourier transform of (4.4) yields:

$$R(\boldsymbol{\xi}) = \vartheta[LAT(\mathbf{V})] \int_{\Upsilon} \Omega_s(\boldsymbol{\nu}) \exp^{j2\pi\boldsymbol{\nu}^T \boldsymbol{\xi}} d\boldsymbol{\nu} \quad (4.5)$$

Let  $R_c(\boldsymbol{\xi})$  denotes the auto-covariance of  $\omega(\boldsymbol{\xi})$ , then by Wiener-Khincine Theorem we have:

$$R_c(\boldsymbol{\xi}) = \int_{\mathbb{R} \times \mathbb{R}^N} \Omega(\boldsymbol{\nu}) \exp^{j2\pi\boldsymbol{\nu}^T \boldsymbol{\xi}} d\boldsymbol{\nu} \quad (4.6)$$

$$= \sum_{\mathbf{r} \in LAT(\mathbf{V}^*)} \int_{\Upsilon} \Omega(\boldsymbol{\nu} + \mathbf{r}) \exp^{j2\pi(\boldsymbol{\nu} + \mathbf{r})^T \boldsymbol{\xi}} d\boldsymbol{\nu} \quad (4.7)$$

By property of the polar lattice,  $\exp^{j2\pi\mathbf{r}^T\boldsymbol{\xi}} = 1$  so that exchanging the order of summation and integration gives:

$$R_c(\boldsymbol{\xi}) = \int_{\Upsilon} \left[ \sum_{\mathbf{r} \in LAT(\mathbf{V}^*)} \Omega(\boldsymbol{\nu} + \mathbf{r}) \right] \exp^{j2\pi\boldsymbol{\nu}^T\boldsymbol{\xi}} d\boldsymbol{\nu}$$

The auto-covariance of the sampled signal,  $R(\boldsymbol{\xi}) = R_c(\boldsymbol{\xi})$ ,  $\boldsymbol{\xi} \in LAT(\mathbf{V})$  so that from (4.5) and (4.6) we have:

$$\Omega_s(\boldsymbol{\nu}) = \frac{1}{\vartheta[LAT(\mathbf{V})]} \sum_{\mathbf{r} \in LAT(\mathbf{V}^*)} \Omega(\boldsymbol{\nu} + \mathbf{r})$$

which gives the desired result. ■

This is the key property of the lattice theoretic framework that makes it so attractive for the optimal  $TS(1)$  sampling design. It enables ease of transforming the temporal-spatial frequency domain requirements to the temporal-spatial domain and vice-versa.

**Property 4.2** *If  $LAT(\widehat{\mathbf{V}}) = LAT(\mathbf{V})$  then there exists a unimodular matrix (i.e. an integer matrix  $\mathbf{Q}$  with  $\det(\mathbf{Q}) = \pm 1$ ) such that  $\widehat{\mathbf{V}} = \mathbf{V}\mathbf{Q}$ . The group of all  $(1+N) \times (1+N)$  unimodular matrices is denoted by  $\mathcal{GL}_n(\mathbb{Z})$ .*

**Proof:**

- The column  $\mathbf{v}_i$  of  $\mathbf{V}$  are points belonging to the lattice  $LAT(\mathbf{V})$ . Since  $\widehat{\mathbf{V}}$  is also a basis, we can express these columns as  $\widehat{\mathbf{V}}\mathbf{z}_i$  for integer  $\mathbf{z}_i$  so we have:

$$\mathbf{V} = \widehat{\mathbf{V}} \underbrace{[ \mathbf{z}_1 \quad \mathbf{z}_2 \quad \dots \quad \mathbf{z}_{1+N} ]}_{\mathbf{R}}$$

- By the same argument we can express each column of  $\widehat{\mathbf{V}}$  as  $\mathbf{V}\mathbf{w}_i$  for integer  $\mathbf{w}_i$  that is:

$$\widehat{\mathbf{V}} = \mathbf{V} \underbrace{[ \mathbf{w}_1 \quad \mathbf{w}_2 \quad \dots \quad \mathbf{w}_{1+N} ]}_{\mathbf{Q}}$$

- Combining the two expressions we obtain:

$$\mathbf{V} = \mathbf{V}\mathbf{Q}\mathbf{R}$$

- Since  $\mathbf{V}$  is nonsingular, this implies that  $\mathbf{Q}\mathbf{R} = \mathbf{I}$  so that:

$$\det(\mathbf{Q})\det(\mathbf{R}) = 1$$



- Since  $\mathbf{Q}$  and  $\mathbf{R}$  are integer matrices, their determinants are integers such that the above equation implies  $\det(\mathbf{Q}) = \pm 1$ .

■

This property means that the lattice representation is not unique. For a given generator matrix  $\mathbf{V} \in \mathbb{R}^{(1+N) \times (1+N)}$ , all lattices  $LAT(\widehat{\mathbf{V}})$  where  $\widehat{\mathbf{V}} = \mathbf{V}\mathbf{Q}$ ,  $\mathbf{Q} \in \mathcal{GL}_n(\mathbb{Z})$  give the same point lattice as  $LAT(\mathbf{V})$ . For example, the generator matrix  $\mathbf{V} = \begin{pmatrix} 1 & 4 \\ 4 & 1 \end{pmatrix}$  and  $\widehat{\mathbf{V}} = \begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}$  forms the same lattice points i.e.  $LAT(\mathbf{V}) = LAT(\widehat{\mathbf{V}})$  since  $\mathbf{V}$  and  $\widehat{\mathbf{V}}$  are related by:

$$\widehat{\mathbf{V}} = \mathbf{V} \underbrace{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}}_{\mathbf{Q} \in \mathbb{Z}^{2 \times 2}}$$

with  $\det(\mathbf{Q}) = 1$  and hence  $\mathbf{Q}$  is a unimodular matrix.

**Property 4.3** *A lattice  $LAT(\mathbf{V})$  is said to be a sublattice of  $LAT(\mathbf{Y})$  if every points of  $LAT(\mathbf{V})$  belongs to  $LAT(\mathbf{Y})$  i.e.  $LAT(\mathbf{V}) \subset LAT(\mathbf{Y})$ . The generator matrices  $\mathbf{Y}$  and  $\mathbf{V}$  are related through the integer decimation/interpolation matrix  $\mathbf{V}_z$  by  $\mathbf{V} = \mathbf{Y}\mathbf{V}_z$ .*

Proof of this property can be found in [50]. This property is useful for deducing a useful relation to characterize  $LAT(\mathbf{V})$  on the temporal-spatial grid points given by (3.14).

**Property 4.4** *The volume of each of the lattice cell,  $\vartheta[LAT(\mathbf{V})]$  is given by the determinant of the generator matrix i.e.  $\vartheta[LAT(\mathbf{V})] = |\det(\mathbf{V})|$  as shown by the shaded area in Figure 4.1.*

By this property, the volume of each lattice cell,  $\vartheta[LAT(\mathbf{V})]$  is unique and independent of the particular choice of the basis. Its reciprocal i.e.  $1/\vartheta[LAT(\mathbf{V})]$  gives the density of the lattice. For the case of  $TS(1)$  sampling lattice, the non-aliasing condition imposes a minimum bound on the packing density given by  $1/\vartheta[LAT(\mathbf{V}^*)]$  where  $\mathbf{V}^*$  gives the polar generator matrix.

**Property 4.5** *If  $\xi_1, \xi_2 \in LAT(\mathbf{V})$  then  $\xi_1 - \xi_2 \in LAT(\mathbf{V})$ .*

In addition to Property 4.1 to Property 4.5, throughout this dissertation *lattice packing* or *packing* will be used to describe the non-aliasing of the ellipsoidal spectral support,  $\Theta$  and

also the non overlapping of the spatial support,  $\mathcal{S}$ . The former is an important property that guarantees the non-aliasing condition. The latter is an important property to fulfill to define  $TS(1)$  sampling lattices.

**Definition 4.1 (*lattice packing*)** A lattice  $LAT(\mathbf{V}) \subset \mathbb{R}^{(1+N)}$  is said to pack a body  $\mathcal{D} \subset \mathbb{R}^{(1+N)}$ , if

$$(\mathcal{D} + \boldsymbol{\xi}_1) \cap (\mathcal{D} + \boldsymbol{\xi}_2) = \emptyset, \forall \boldsymbol{\xi}_1, \boldsymbol{\xi}_2 \in LAT(\mathbf{V}), \boldsymbol{\xi}_1 \neq \boldsymbol{\xi}_2 \quad (4.8)$$

In other words,  $\mathcal{D}$  translates to  $\boldsymbol{\xi} \in LAT(\mathbf{V})$  do not intersect each other. Figure 3.8(b) shows an example of the *lattice packing* for the ellipses,  $\Theta$ .

The use of the lattice theoretic framework for the general multidimensional sampling problem was first proposed by Peterson and Middleton [17]. They were the first to demonstrate Property 4.1 where a multidimensional signal sampled on a lattice  $LAT(\mathbf{V})$  results in the replication of the original frequency content of the signal on a lattice  $LAT(\mathbf{V}^{-T})$  in the frequency domain. This enables the well known Whittaker-Kotel'nikov-Shannon Sampling Theorem [48] for frequency-bandlimited functions of time to be extended to functions of multidimensional arguments. Therefore a multidimensional signal whose spectrum is restricted to a bandlimited region, may be reconstructed from its samples taken over a periodic lattice having suitably small repetition vectors. Mersereau and Speake [51] and Dubois [52] further developed the lattice framework, in particular showing that most common signal processing operations, such as convolution, recursive and nonrecursive filtering, decimation and interpolation, Fourier analysis, and discrete Fourier analysis can be performed on any periodic sampling lattice. These works set the stage for evaluating generalized sampling systems (e.g. hexagonal sampling, field-quincunx sampling, line-quincunx sampling etc.) of multidimensional signals. This is important because the most efficient sampling systems of multidimensional signals depend on signals spectrum and dimensionality. For example, it is shown in [51] that hexagonal sampling and its higher dimensional generalizations resulted in lower sampling density and more efficient signal processing algorithms than common rectangular sampling for isotropically band-limited signals.

However traditionally, discrete processing of multidimensional signals placed no distinction between the time and spatial dimension on the lattice. For example, in the sampling of time-varying imagery, samples from the entire two-dimensional spatial lattice are available at each processing instance, and therefore we are not constrained to treat the time and spatial dimension differently. In time-sequential sampling,  $n$  samples are taken one at a time according to a specific periodic ordering. This means that only certain special lattices

are time-sequential sampling. Willis and Bresler [13, 14] gave a rather complete analysis of  $TS(1)$  sampling lattice (i.e.  $n = 1$  spatial samples at each time instance) and laid the foundation for the design of an optimal  $TS(1)$  sampling lattice. This work presented a unifying theory linking  $TS(1)$  sampling with classical multidimensional sampling in the lattice theoretic framework. Specifically, by making use of Hermite Normal Form, the condition for a lattice to be a  $TS(1)$  sampling lattice with respect to an arbitrary spatial support is given in this work. In this next section, the Hermite Normal Form and its use in the sampling framework as given in Chapter 3 for defining the  $TS(n)$  sampling is presented.

## 4.2 $TS(n)$ Sampling Lattices

In this section,  $TS(n)$  sampling for a  $1 + N$ -dimensional temporal-spatial signal is described on the lattice framework. Let  $\boldsymbol{\xi} = [t_k, \xi_1, \dots, \xi_N] \in LAT(\mathbf{V})$  where  $t_k = kT$  is the temporal coordinate and  $[\xi_1, \dots, \xi_N] =: \boldsymbol{\xi}_s$  are the  $N$ -dimensional spatial (or generalized-tonal) coordinates to be sampled at time  $k$ . Also let,

$$\mathcal{T}(\mathbf{V}) = \{\tau | [\tau, \boldsymbol{\xi}_s]^T \in LAT(\mathbf{V})\} \quad (4.9)$$

which defines the set of all sampling times.

**Definition 4.2 ( $TS(n)$  sampling lattice)**  *$LAT(\mathbf{V})$  is a  $TS(n)$  sampling lattice if for each  $t_k \in \mathcal{T}(\mathbf{V})$  there are at most  $n$  unique points  $\boldsymbol{\xi} = [t_k, \boldsymbol{\xi}_s]^T \in LAT(\mathbf{V})$  such that  $\boldsymbol{\xi}_s^T \in \mathcal{S} \subset \mathbb{R}^N$  where  $\mathcal{S}$  is the spatial support.*

However, not all lattices are  $TS(n)$  sampling. For example, the lexicographic sequence as shown in Figure 3.4 for the case of the TRC *can* be described by a lattice but the bit-reversed sequence shown in the same figure *cannot* be described by a lattice. Hence we need to characterize all the  $TS(n)$  sampling lattices i.e. we would like to find necessary and sufficient conditions on the generator matrix  $\mathbf{V}$  such that  $LAT(\mathbf{V})$  is  $TS(n)$  sampling lattice. Let  $\mathbf{V} \in \mathbb{R}^{(1+N) \times (1+N)}$  be a non-singular matrix. The lattice generated by  $\mathbf{V}$ ,  $LAT(\mathbf{V}) \subset \mathbb{R}^{(1+N)}$ , is a set of points given by (4.1). We start by noting that in  $TS(n)$  sampling,  $\boldsymbol{\omega}(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$  is sampled on a set of temporal-spatial grid points given by (3.14). In the lattice framework,  $\mathcal{G}$  in (3.14) is  $LAT(\mathbf{G})$  where

$$\mathbf{G} = \begin{pmatrix} T & 0 \\ 0 & \mathbf{G}_1 \end{pmatrix} \in \mathbb{R}^{(1+N) \times (1+N)} \quad (4.10)$$

with  $\mathbf{G}_1 = \text{diag}(\mathbf{A} \cdot / \mathbf{M}) \in \mathbb{R}^{N \times N}$ .  $\mathbf{A} = [A_1, \dots, A_N]^T$  is define as the vector of the  $N$ -dimensional maximum spatial range and  $\mathbf{M} = [M_1, \dots, M_N]^T$  is the vector of  $N$ -dimensional

number of points that uniformly discretize the spatial range (see Figure 3.3 for a pictorial description of these parameters). Hence, the notation  $\mathbf{A} \cdot / \mathbf{M} = [A_1/M_1, \dots, A_N/M_N]$ . By this construction, the  $TS(n)$  sampling lattice  $LAT(\mathbf{V})$  is a sublattice of  $LAT(\mathbf{G})$ . By Property 4.3 the matrices  $\mathbf{V}$  and  $\mathbf{G}$  is then related by

$$\mathbf{V} = \mathbf{G}\mathbf{V}_z \in \mathbb{R}^{(1+N) \times (1+N)} \quad (4.11)$$

where  $\mathbf{V}_z \in \mathbb{Z}^{(1+N) \times (1+N)}$  is an integer decimation/interpolation matrix. In this formulation,  $LAT(\mathbf{V}_z)$  gives the  $TS(n)$  sampling integer sequence i.e.  $\boldsymbol{\alpha}(k)$  given by (3.18), on the grid,  $LAT(\mathbf{G})$ . Additionally, the factorized form of (4.11) allows for the use of the canonical Hermite Normal Form (HNF) to define unique lattices. This is important because without the HNF form, the basis  $\mathbf{V}$  in  $LAT(\mathbf{V})$  is not unique (by Property 4.2). The HNF form which gives a unique representation for any lattice is given as follows:

**Theorem 4.2 (HNF)** [53] *A  $(1 + N) \times (1 + N)$  nonsingular integer matrix  $\mathbf{V}_z$  can be uniquely decomposed as*

$$\mathbf{V}_z = \mathbf{H}_{LTM}\mathbf{Q} \in \mathbb{Z}^{(1+N) \times (1+N)} \quad (4.12)$$

where  $\mathbf{Q}$  is a unimodular ( $\det(\mathbf{Q}) = \pm 1$ )  $(1 + N) \times (1 + N)$  integer matrix, and

$$\mathbf{H}_{LTM} = \begin{pmatrix} {}^\ell h_{11} & 0 & \dots & 0 \\ {}^\ell h_{21} & {}^\ell h_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ {}^\ell h_{1+N,1} & {}^\ell h_{1+N,2} & \dots & {}^\ell h_{1+N,1+N} \end{pmatrix} \quad (4.13)$$

is a lower triangular (column-reduced) integer matrix such that:

- (i)  ${}^\ell h_{ij} = 0$ , for  $i < j$
- (ii)  ${}^\ell h_{ii} > 0 \forall i$ , and
- (iii)  ${}^\ell h_{ij} \leq 0$  and  $|{}^\ell h_{ij}| < {}^\ell h_{ii}$ , for  $i > j$

Therefore, by this theorem, two integer generator matrices that generate the same lattice,  $LAT(\mathbf{V}_{z1}) = LAT(\mathbf{V}_{z2})$  have the same HNF so that  $\mathbf{V}_{z1} = \mathbf{H}_{LTM}\mathbf{Q}_1$  and  $\mathbf{V}_{z2} = \mathbf{H}_{LTM}\mathbf{Q}_2$ . Moreover,  $\mathbf{V}_{z1} = \mathbf{V}_{z2}\mathbf{Q}_{12}$  where  $\mathbf{Q}_{12} = \mathbf{Q}_2\mathbf{Q}_1^{-1}$  is a unimodular matrix.

To obtain the HNF form,  $\mathbf{H}_{LTM} \in \mathbb{Z}^{(1+N) \times (1+N)}$  given any integer matrix,  $\mathbf{V}_z \in \mathbb{Z}^{(1+N) \times (1+N)}$ , the column-reduced algorithm given in Appendix C.2 is used.

The lower-triangular HNF form gives useful geometric insights of the lattice that can be used to define the  $TS(n)$  sampling lattices [13]. From (4.12), we have:

$$\mathbf{V} = \mathbf{G}\mathbf{V}_z = \mathbf{G}\mathbf{H}_{LTM}\mathbf{Q}$$

where  $\mathbf{Q}$  is a unimodular matrix. By lattice Property 4.2,  $LAT(\mathbf{V}) = LAT(\mathbf{G}\mathbf{H}_{LTM})$ . Let  $\boldsymbol{\xi} \in LAT(\mathbf{G}\mathbf{H}_{LTM})$ , where  $\boldsymbol{\xi} := [t_k, \boldsymbol{\xi}_s]^T = \mathbf{G}\mathbf{H}_{LTM}^\ell \mathbf{z}$  for some  ${}^\ell \mathbf{z} = \begin{pmatrix} {}^\ell z_1 \\ \vdots \\ {}^\ell \mathbf{z}_s \end{pmatrix} \in \mathbb{Z}^{(1+N)}$ . Expanding this we have (4.14) and (4.15).

$$t_k = T^\ell h_{11} {}^\ell z_1 \quad (4.14)$$

$$\begin{pmatrix} \xi_{s,1} \\ \xi_{s,2} \\ \vdots \\ \xi_{s,N} \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{A_1}{} {}^\ell h_{21} \\ \frac{A_2}{} {}^\ell h_{31} \\ \vdots \\ \frac{A_N}{} {}^\ell h_{N+1,1} \end{pmatrix}}_{\mathbf{s}} {}^\ell z_1 + \underbrace{\begin{pmatrix} \frac{A_1}{} {}^\ell h_{22} & 0 & \dots & 0 \\ \frac{A_2}{} {}^\ell h_{32} & \frac{A_2}{} {}^\ell h_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_N}{} {}^\ell h_{1+N,2} & \frac{A_N}{} {}^\ell h_{1+N,3} & \dots & \frac{A_N}{} {}^\ell h_{1+N,1+N} \end{pmatrix}}_{\nabla} {}^\ell \mathbf{z}_s \quad (4.15)$$

where  ${}^\ell \mathbf{z}_s = [{}^\ell z_2, \dots, {}^\ell z_{1+N}]^T \in \mathbb{Z}^N$ . From (4.14), since  ${}^\ell z_1 \in \mathbb{Z}$  and  $t_k = kT$  means that  ${}^\ell h_{11} = 1$  with  ${}^\ell z_1 = k$ . Collectively, we can rewrite (4.14) and (4.15) compactly as

$$\begin{pmatrix} t_k \\ \boldsymbol{\xi}_s^T \end{pmatrix} = \underbrace{\begin{pmatrix} T & 0 \\ \mathbf{s} & \nabla \end{pmatrix}}_{\mathbf{G}\mathbf{H}_{LTM}} \begin{pmatrix} k \\ {}^\ell \mathbf{z}_s \end{pmatrix} \quad (4.16)$$

Hence, by condition on the lower triangular HNF,  $\mathbf{H}_{LTM} \in \mathbb{Z}^{(1+N) \times (1+N)}$  given by Theorem 4.2, the set of all unique lattices on the grid  $LAT(\mathbf{G})$  is given by:

$$\mathcal{L}(\mathbf{G}) = \left\{ LAT(\mathbf{G}\mathbf{H}_{LTM}) \left| \begin{array}{l} \mathbf{H}_{LTM} \in \mathbb{Z}^{(1+N) \times (1+N)} \text{ by (4.13) takes the values of:} \\ (i) {}^\ell h_{ij} = 0, \text{ for } i < j, \\ (ii) {}^\ell h_{11} = 1 \\ (iii) {}^\ell h_{ii} > 0, \forall i \in \{2, \dots, 1+N\} \\ (iv) {}^\ell h_{ij} \leq 0 \text{ and } |{}^\ell h_{ij}| < {}^\ell h_{ii}, i > j \end{array} \right. \right\} \quad (4.17)$$

From (4.16), the spatial lattice points at each time step  $kT$  are given by:

$$\boldsymbol{\xi}_s^T = k\mathbf{s} + \nabla {}^\ell \mathbf{z}_s \in k\mathbf{s} + LAT(\nabla)$$

Hence by the definition of  $TS(n)$  (Definition 4.2) the  $TS(n)$  sampling lattice with respect to a known spatial support,  $\mathcal{S}$  can now be described by the following lemma.

**Lemma 4.1** *The lattice  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G})$  gives a  $TS(n)$  sampling lattice on the grid  $LAT(\mathbf{G})$  with respect to the spatial support  $\mathcal{S}$ , if and only if*

$$|\mathcal{S} \cap (k\mathbf{s} + LAT(\nabla))| \leq n, \quad \forall k \in \mathbb{Z}^+$$

where  $|\cdot|$  gives the cardinality of the finite set.

**Proof:** Since the spatial lattice points at each time step  $kT$  are given by  $k\mathbf{s} + LAT(\nabla)$ , then by Definition 4.2, we arrived at the desired result.  $\blacksquare$

Lemma 4.1 gives an insightful relation of the HNF form and the  $TS(n)$  sampling lattice. While this condition can be verified for a given lattice, it is difficult in general, to specify explicitly the set of  $\mathbf{H}_{LTM} \in \mathbb{Z}^{(1+N) \times (1+N)}$  matrices such that this is true, which is needed for efficient design. For  $TS(1)$  sampling, however, a more useful condition can be defined to enable parameterizations of  $TS(1)$  sampling lattices possible as will be demonstrated in the next section.

### 4.3 Parameterizing $TS(1)$ Sampling Lattices

In this section, the space of all  $TS(1)$  sampling lattices with respect to the rectangular spatial support,  $\mathcal{S} \subset \mathbb{R}^N$  on grid  $LAT(\mathbf{G})$  will be parameterized by the inter-sampling time  $T \in \mathbb{R}$ , the number of spatial discretization at each dimension  $M_i \in \mathbb{Z}^+$  and the spatial range  $A_i \in \mathbb{R}^+$ . Let  $\mathcal{E}(1, \mathcal{S})$  denote the set of all  $TS(1)$  sampling (including non-lattice  $TS(1)$  sampling) with respect to the spatial domain, i.e.

$$\mathcal{E}(1, \mathcal{S}) = \{(\alpha_1(k), T, \mathbf{A}, \mathbf{M}) \mid (\alpha_1(k), T, \mathbf{A}, \mathbf{M}) \text{ gives a } TS(1) \text{ sampling}\} \quad (4.18)$$

The quadruple  $(\alpha_1(k), T, \mathbf{A}, \mathbf{M})$  is described in section 3.2.1. By the lattice set given in (4.17), the set of all  $TS(1)$  sampling lattices with respect to  $\mathcal{S} \subset \mathbb{R}^N$  on grid  $LAT(\mathbf{G})$  is therefore given by  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ .

A necessary and sufficient condition to define the set  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  will be given in Theorem 4.3. This condition looks at how the translates of the spatial domain on the lattice points,  $LAT(\nabla)$  i.e.  $\mathcal{S} + \mathbf{b}, \mathbf{b} \in \mathcal{S}, \mathbf{b} \neq \mathbf{0}$  overlaps the original spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$ . If no overlaps occurs, then only one spatial point i.e.  $\mathbf{0} \in LAT(\nabla)$  point is in  $\mathcal{S}$  at all time-step

$k \in \mathbb{Z}^+$ . If the spatial domain overlaps i.e.  $\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \neq \emptyset$  for some  $\mathbf{b} \in LAT(\nabla)$ ,  $\mathbf{b} \neq \mathbf{0}$ , then the *possibility* exists such that at some time-step, there are more than one point in  $\mathcal{S}$ . In this case  $LAT(\mathbf{GH}_{LTM})$  is not a  $TS(1)$  sampling lattice. To negate this possibility, we need the collection of all spatial sampling points i.e.  $\bigcup_{k \in \mathbb{Z}^+} (LAT(\nabla) + ks)$  and ensure that none of these points are in the overlap region.

The collection of all the spatial sampling points can be described by the upper triangular HNF realization of the generator matrix for the lattice,  $LAT(\mathbf{V}) = LAT(\mathbf{GH}_{LTM})$ . To arrive at this realization, consider first the inverse transpose of the generator matrix  $\mathbf{GH}_{LTM} \in \mathbb{R}^{(1+N) \times (1+N)}$ ,

$$(\mathbf{GH}_{LTM})^{-T} = \frac{\mathbf{G}^{-T}}{\det(\mathbf{H}_{LTM})} \bar{\mathbf{V}}_z \quad (4.19)$$

where  $\bar{\mathbf{V}}_z := \det(\mathbf{H}_{LTM}) \mathbf{H}_{LTM}^{-T} \in \mathbb{Z}^{(1+N) \times (1+N)}$  is now an upper triangular integer matrix. By Theorem 4.2,  $\bar{\mathbf{V}}_z \in \mathbb{Z}^{(1+N) \times (1+N)}$  can be written as  $\bar{\mathbf{V}}_z = \bar{\mathbf{H}}_{LTM} \bar{\mathbf{Q}}$ , where  $\bar{\mathbf{Q}} \in \mathbb{Z}^{(1+N) \times (1+N)}$  is a unimodular matrix and  $\bar{\mathbf{H}}_{LTM} \in \mathbb{Z}^{(1+N) \times (1+N)}$  is the lower triangular HNF form. Hence, by inverse transpose equation (4.19), we have:

$$\mathbf{GH}_{LTM} = \underbrace{\det(\mathbf{H}_{LTM}) \mathbf{G} \bar{\mathbf{H}}_{LTM}^{-T}}_{\mathbf{V}_{UTM}} \bar{\mathbf{Q}}^{-T} \quad (4.20)$$

where  $\bar{\mathbf{Q}}^{-T}$  is still unimodular and  $\mathbf{V}_{UTM} \in \mathbb{R}^{(1+N) \times (1+N)}$  is an upper triangular form with diagonal that is greater than 0. By lattice Property 4.2,  $LAT(\mathbf{GH}_{LTM}) = LAT(\mathbf{V}_{UTM})$ . Let  $\zeta \in LAT(\mathbf{V}_{UTM})$ , where  $\zeta := [\tau, \zeta_s]^T = \mathbf{V}_{UTM}^u \mathbf{z}$  for some  ${}^u \mathbf{z} = \begin{pmatrix} {}^u z_1 \\ {}^u \mathbf{z}_s \end{pmatrix} \in \mathbb{Z}^{(1+N)}$ . Then,

$$\begin{pmatrix} \tau \\ \zeta_s^T \end{pmatrix} = \underbrace{\begin{pmatrix} q & \mathbf{r} \\ \mathbf{0} & \Delta \end{pmatrix}}_{\mathbf{V}_{UTM}} \begin{pmatrix} {}^u z_1 \\ {}^u \mathbf{z}_s \end{pmatrix} \quad (4.21)$$

In this form, each spatial lattice points,  $\zeta_s$  are specified at time instance,  $\tau \in \mathbb{R}$  and these points do not depend on the time-index  ${}^u z_1 = k \in \mathbb{Z}$ . Since,  $LAT(\mathbf{GH}_{LTM}) = LAT(\mathbf{V}_{UTM})$  and comparing (4.16) and (4.21), the *collection* of the spatial lattice points using the lower-triangular HNF and the *collection* of the spatial lattice points using the upper-triangular form can be related by:

$$LAT(\Delta) = \bigcup_{k \in \mathbb{Z}} LAT(\nabla) + ks \quad (4.22)$$

The necessary and sufficient condition that defines the set  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  can now be given

by the following theorem.

**Theorem 4.3** [13] *For a given spatial support, the lattice  $LAT(\mathbf{GH}_{LTM})$  belongs to the set  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  where  $\mathbf{GH}_{LTM}$  is of the form (4.16) (i.e. it is a  $TS(1)$  sampling lattice) if and only if*

$$\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta) = \emptyset, \quad \forall \mathbf{b} \in LAT(\nabla) \subset \mathbb{R}^N, \mathbf{b} \neq \mathbf{0} \quad (4.23)$$

**Proof:** ( $\Rightarrow$ )

- Suppose  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  but (4.23) is violated. Hence  $\exists \mathbf{b} \in LAT(\nabla)$ ,  $\mathbf{b} \neq \mathbf{0}$  and  $\xi_1 \in LAT(\Delta)$  such that  $\xi_1 \in \mathcal{S}$  and  $\xi_1 \in \mathcal{S} + \mathbf{b}$ .
- Let  $\xi_2 := \xi_1 - \mathbf{b}$ . Hence  $\xi_1, \xi_2 \in \mathcal{S}$ . Since  $\mathbf{b} \neq \mathbf{0}$ ,  $\xi_1 \neq \xi_2$ .
- Since  $LAT(\Delta) = \bigcup_{k \in \mathbb{Z}} LAT(\nabla) + k\mathbf{s}$ ,  $\xi_1 \in LAT(\nabla) + \bar{k}\mathbf{s}$  for some  $\bar{k} \in \mathbb{Z}$ .
- Since  $\mathbf{b} \in LAT(\nabla)$ , by lattice Property 4.5,  $\xi_1, \xi_2 \in LAT(\nabla) + \bar{k}\mathbf{s}$ .
- Hence, there exists  $\xi_1, \xi_2 \in \mathcal{S} \cap (LAT(\nabla) + \bar{k}\mathbf{s})$ , for some  $\bar{k}$  and  $\xi_1 \neq \xi_2$ . By Lemma 4.1,  $LAT(\mathbf{GH}_{LTM})$  is not a  $TS(1)$  sampling lattice which contradicts our assumption.

( $\Leftarrow$ )

- Suppose (4.23) is true but  $LAT(\mathbf{GH}_{LTM})$  does not belong to the set of  $TS(1)$  sampling lattice,  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ . Then by Lemma 4.1 at some time step  $\bar{k} \in \mathbb{Z}$  there exists at least two spatial points  $\xi_1, \xi_2 \in \mathcal{S}, \xi_1 \neq \xi_2$  such that  $\xi_1, \xi_2 \in LAT(\nabla) + \bar{k}\mathbf{s}$ .
- Let  $\tilde{\xi}_1 = \xi_1 - \bar{k}\mathbf{s} \in LAT(\nabla)$  and  $\tilde{\xi}_2 = \xi_2 - \bar{k}\mathbf{s} \in LAT(\nabla)$ . Also let  $\mathbf{b} = \xi_1 - \xi_2 = \tilde{\xi}_1 - \tilde{\xi}_2$ . Now  $\tilde{\xi}_1, \tilde{\xi}_2 \in LAT(\nabla)$  and by lattice Property 4.5,  $\mathbf{b} \in LAT(\nabla)$ .
- Now  $\xi_1 \in \mathcal{S}, \xi_2 = \xi_1 - \mathbf{b} \in \mathcal{S}$  and also, since  $LAT(\nabla) + \bar{k}\mathbf{s} \subset LAT(\Delta)$ ,  $\xi_1 \in LAT(\Delta)$ .
- Hence  $\xi_1 \in \mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta)$ . This contradicts the assumption that (4.23) is true.

■



**Remark 4.1** Note that in Theorem 4.3,  $\mathcal{S}$  does not necessarily be a rectangular spatial support. This restriction is not used in the proof of this theorem. In fact this theorem is applicable to any bounded closed set with nonempty interior with its boundary having a Lesbegue measure zero [13].

Theorem 4.3 characterized  $TS(1)$  sampling lattices by checking the condition given by (4.23). As will be shown next, by this condition and by restricting the spatial domain to be a rectangular space allows us to parameterize the set of all possible  $TS(1)$  sampling lattices on a given grid,  $LAT(\mathbf{G})$ .

To define this set we first restrict the spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$  to a rectangular space as given by (3.1) i.e.

$$\mathcal{S} = \prod_{i=1}^N [0, A_i)$$

By this restriction, we derive conditions that directly restrict the values of the  $\mathbf{H}_{LTM}$  matrix can take for a given grid,  $LAT(\mathbf{G})$  such that  $LAT(\mathbf{G}\mathbf{H}_{LTM})$  is a  $TS(1)$  sampling lattice<sup>1</sup>. The spatial domain restriction is not deemed severe because we can often find a bijective function,  $\mathcal{X} : \mathcal{D} \rightarrow \mathcal{S}$ , where  $\mathcal{D}$  is a non-rectangular spatial domain.

The following defines a subset of  $TS(1)$  sampling lattices. It gives the required restriction on the elements of the  $\mathbf{H}_{LTM}$  matrix for a given grid,  $LAT(\mathbf{G})$  such that  $LAT(\mathbf{G}\mathbf{H}_{LTM})$  is a  $TS(1)$  sampling lattice.

$$\mathcal{LE}_s(\mathbf{G}) = \left\{ LAT(\mathbf{G}\mathbf{H}_{LTM}) \left| \begin{array}{l} \mathbf{H}_{LTM} \text{ by (4.13) takes the values of:} \\ (i) \quad {}^\ell h_{ij} = 0, \text{ for } i < j, \\ (ii) \quad {}^\ell h_{11} = 1 \\ (iii) \quad {}^\ell h_{ii} > 0, \forall i \in \{2, \dots, 1+N\} \\ (iv) \quad {}^\ell h_{ij} \leq 0 \text{ and } |{}^\ell h_{ij}| < {}^\ell h_{ii}, i > j \\ (v) \quad {}^\ell h_{ii} \geq M_{i-1}, \forall i \in \{2, \dots, 1+N\} \end{array} \right. \right\} \subset \mathcal{L}(\mathbf{G}) \quad (4.24)$$

Corollary 4.1 justifies  $\mathcal{LE}_s(\mathbf{G})$  to be the set of  $TS(1)$  sampling lattices. Note that elements (i) to (iv) are conditions needed to define unique lattices as given by the set of all unique lattice,  $\mathcal{L}(\mathbf{G})$  in (4.17). Element (v) is the only element specific to the set  $\mathcal{LE}_s(\mathbf{G})$ .

**Corollary 4.1** If  $LAT(\mathbf{G}\mathbf{H}_{LTM}) \in \mathcal{LE}_s(\mathbf{G})$  then  $LAT(\mathbf{G}\mathbf{H}_{LTM})$  is a  $TS(1)$  sampling lattice i.e.  $LAT(\mathbf{G}\mathbf{H}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  where  $\mathcal{S} = \prod_{i=1}^N [0, A_i)$ .

---

<sup>1</sup>Note that given that  $LAT(\mathbf{G}\mathbf{H}_{LTM})$  is a  $TS(1)$  sampling lattice, the  $(T, \mathbf{A}, \mathbf{M})$  parameters are encoded in matrix  $\mathbf{G}$  and the  $TS(1)$  sampling sequence,  $\alpha_1(k)$  is encoded in the integer  $\mathbf{H}_{LTM}$  matrix.

**Proof:**

- Let  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LE}_s(\mathbf{G})$  where  $\mathbf{GH}_{LTM} = \begin{pmatrix} T & 0 \\ \mathbf{s} & \nabla \end{pmatrix}$  from (4.16). Suppose that  $LAT(\mathbf{GH}_{LTM})$  is not a  $TS(1)$  sampling lattice. From Theorem 4.3,  $\exists \mathbf{b} \in LAT(\nabla), \mathbf{b} \neq \mathbf{0}$  such that  $\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta) \neq \emptyset$ .
- Hence  $\exists \mathbf{b} \in LAT(\nabla), \mathbf{b} \neq \mathbf{0}$  and  $\exists \xi_1 \in LAT(\Delta)$  such that  $\xi_1 \in \mathcal{S}$  and  $\xi_1 \in \mathcal{S} + \mathbf{b}$ .
- Since  $\mathcal{S} = \prod_{i=1}^N [0, A_i)$ ,  $\xi_1 \in \mathcal{S}$  and  $\xi_1 - \mathbf{b} \in \mathcal{S}$  implies that

$$-A_i < b_i < A_i, \forall i = \{1, \dots, N\} \quad (4.25)$$

Since  $\mathbf{b} \in LAT(\nabla)$ , from (4.15) we have,

$$\underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}}_{\mathbf{b}} = \underbrace{\begin{pmatrix} \frac{A_1}{M_1} \ell h_{22} & 0 & \dots & 0 \\ \frac{A_2}{M_2} \ell h_{32} & \frac{A_2}{M_2} \ell h_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_N}{M_N} \ell h_{1+N,2} & \frac{A_N}{M_N} \ell h_{1+N,3} & \dots & \frac{A_N}{M_N} \ell h_{1+N,1+N} \end{pmatrix}}_{\nabla} \underbrace{\begin{pmatrix} \ell z_2 \\ \ell z_3 \\ \vdots \\ \ell z_{1+N} \end{pmatrix}}_{\ell \mathbf{z}_s} \quad (4.26)$$

- Since  $\mathbf{b} \neq \mathbf{0}$ , let  $b_i$  be the first non-zero element of  $\mathbf{b}$  i.e.

$$\mathbf{b} = [0, \dots, 0, b_i, b_{i+1}, \dots, b_N]^T \quad (4.27)$$

where  $b_{i+1}, \dots, b_N$  are real value elements that can be zero.

- By HNF theorem (Theorem 4.2),  $\ell h_{j+1,j+1} > 0$ , hence,  $\frac{A_j}{M_j} \ell h_{j+1,j+1} > 0$  ( $A_j \in \mathbb{R}^+$  and  $M_j \in \mathbb{Z}^+$ ). Hence, from (4.26) for  $b_j = 0$ , implies that  $\ell z_2, \dots, \ell z_{j+1}$  are zero.
- Hence, from (4.26), the first non-zero element of  $\mathbf{b}$ ,  $b_i$  given by (4.27) is:

$$b_i = \frac{A_i}{M_i} \ell h_{i+1,i+1} \ell z_{i+1} \neq 0 \quad (4.28)$$

Note that, since  $b_i \neq 0$  and  $\frac{A_i}{M_i} \ell h_{i+1,i+1} > 0$  this implies that  $\ell z_{i+1} \neq 0$  i.e.  $\ell z_{i+1} \in \mathbb{Z}/\{0\}$ .

- From (4.25), and (4.28) we have:

$$-M_i / \ell z_{i+1} < \ell h_{i+1,i+1} < M_i / \ell z_{i+1} \quad (4.29)$$

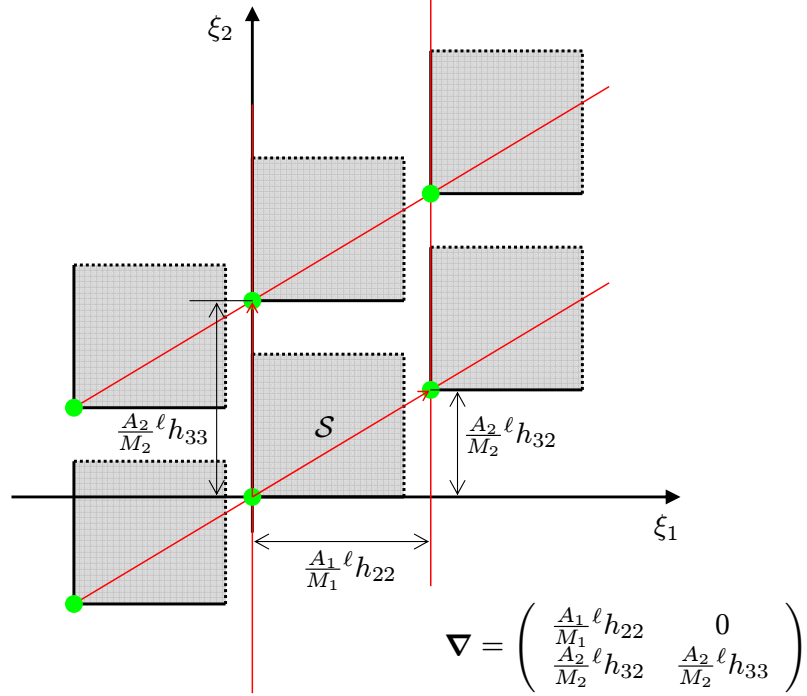


Figure 4.2: The replicas of the spatial support  $\mathcal{S} \subset \mathbb{R}^N$  on  $LAT(\nabla)$ . Note that for  $\ell h_{22} \geq M_1$  and  $\ell h_{33} \geq M_2$ , regardless of what values  $\ell h_{32} \in \mathbb{Z}^+ < \ell h_{33}$  takes, the spatial support will never overlap each other.

Since by HNF theorem,  $\ell h_{i+1,i+1} > 0$ , we can ignore the negative values of  $\ell h_{i+1,i+1}$  in inequality (4.29). Hence we have  $\ell h_{i+1,i+1} < M_i / \ell z_{i+1}$

- Since,  $\ell z_{i+1} \in \mathbb{Z}/\{0\}$ ,  $M_i / \ell z_{i+1} \leq M_i$ . Together this implies that  $\ell h_{i+1,i+1} < M_i$ .
- This implies that there always exists  $i \in \{1, \dots, N\}$  such that  $\ell h_{i+1,i+1} < M_i$ . Hence  $LAT(\mathbf{GH}_{LTM})$  does not belong to the set  $\mathcal{LE}_s(\mathbf{G})$ , contradicting our assumption. ■

Corollary 4.1 gives the sufficient condition for a lattice to be a  $TS(1)$  sampling lattice. Hence by Corollary 4.1 all lattices belonging to the set  $\mathcal{LE}_s(\mathbf{G})$  is a  $TS(1)$  sampling lattice i.e.  $\mathcal{LE}_s(\mathbf{G}) \subseteq \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ . Note that all lattices,  $LAT(\mathbf{GH}_{LTM})$  in the set  $\mathcal{LE}_s(\mathbf{G})$  have diagonal of  $\nabla \in \mathbb{R}^{N \times N}$  that is greater or equal to the spatial range of the rectangular spatial support,  $\mathcal{S} \subset \mathbb{R}^N$  i.e. from (4.24) we have that in this set  $\ell h_{ii} \geq M_{i-1}, \forall i \in \{2, \dots, 1+N\}$ . With  $i \rightarrow i+1$  we have:

$$\frac{A_i}{M_i} \ell h_{i+1,i+1} \geq A_i, \forall i \in \{1, \dots, N\} \quad (4.30)$$

The significance of this inequality can be pictorially shown in Figure 4.2 for  $1 + N = 3$  dimensional temporal-spatial signal where we observed that the replicas of the spatial support,  $\mathcal{S} \subset \mathbb{R}^N$  on  $LAT(\nabla)$  will never overlap each other when inequality (4.30) is obeyed. Theorem 4.3 is fulfilled and all lattices,  $LAT(\mathbf{GH}_{LTM})$  in the set of  $\mathcal{LE}_s(\mathbf{G})$  are  $TS(1)$  sampling lattices.

To formulate the necessary condition, the following defines the superset of all  $TS(1)$  sampling lattices.

$$\mathcal{LE}_n(\mathbf{G}) = \left\{ LAT(\mathbf{GH}_{LTM}) \left| \begin{array}{l} \mathbf{H}_{LTM} \text{ by (4.13) takes the values of:} \\ (i) \quad \ell h_{ij} = 0, \text{ for } i < j, \\ (ii) \quad \ell h_{11} = 1 \\ (iii) \quad \ell h_{ii} > 0, \forall i \in \{2, \dots, 1 + N\} \\ (iv) \quad \ell h_{ij} \leq 0 \text{ and } |\ell h_{ij}| < \ell h_{ii}, i > j \\ (v) \quad \ell h_{1+N, 1+N} \geq M_N \end{array} \right. \right\} \subset \mathcal{L}(\mathbf{G}) \quad (4.31)$$

Note that elements (i) to (iv) are conditions needed to define unique lattices as given by the set of all unique lattice,  $\mathcal{L}(\mathbf{G})$  in (4.17). Element (v) is the only element specific to the set  $\mathcal{LE}_n(\mathbf{G})$ .

**Corollary 4.2** *If  $LAT(\mathbf{GH}_{LTM})$  is a  $TS(1)$  sampling lattice with respect to the spatial domain,  $\mathcal{S} = \prod_{i=1}^N [0, A_i]$  on grid,  $LAT(\mathbf{G})$  i.e.  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  then  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LE}_n(\mathbf{G})$*

**Proof:**

- Let  $LAT(\mathbf{GH}_{LTM})$  where  $\mathbf{GH}_{LTM} = \begin{pmatrix} T & 0 \\ \mathbf{s} & \nabla \end{pmatrix}$  from (4.16) be a  $TS(1)$  sampling lattice i.e.  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ . Suppose  $LAT(\mathbf{GH}_{LTM})$  does not belong to the set of  $\mathcal{LE}_n(\mathbf{G})$ . This means that the  $\mathbf{H}_{LTM}$  matrix given by (4.13) takes the values of:

- (i)  $\ell h_{ij} = 0, \text{ for } i < j$
- (ii)  $\ell h_{11} = 1$
- (iii)  $\ell h_{ii} > 0, \forall i \in \{2, \dots, 1 + N\}$
- (iv)  $\ell h_{ij} \leq 0 \text{ and } |\ell h_{ij}| < \ell h_{ii}, i > j$
- (v)  $\ell h_{1+N, 1+N} < M_N$

- Now, pick,

$$\mathbf{b}^* = \underbrace{\begin{pmatrix} \frac{A_1}{M_1} \ell h_{22} & 0 & \dots & 0 \\ \frac{A_2}{M_2} \ell h_{32} & \frac{A_2}{M_2} \ell h_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_N}{M_N} \ell h_{1+N,2} & \frac{A_N}{M_N} \ell h_{1+N,3} & \dots & \frac{A_N}{M_N} \ell h_{1+N,1+N} \end{pmatrix}}_{\nabla} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \frac{A_N \ell h_{1+N,1+N}}{M_N} \end{pmatrix} \quad (4.32)$$

Since  $0 < \ell h_{1+N,1+N} < M_N$ , we have

$$0 < \frac{A_N \ell h_{1+N,1+N}}{M_N} < A_N$$

By (4.32),  $\mathbf{b}^* \in \mathcal{S} = \prod_{i=1}^N [0, A_i]$ .

- This implies that  $\mathbf{0} \in \mathcal{S} + (-\mathbf{b}^*)$ . Also  $\mathbf{0} \in \mathcal{S}$ .
- Let  $\mathbf{b}^\dagger = -\mathbf{b}^*$ . By Property 4.5, since  $\mathbf{b}^\dagger = \mathbf{b}^* - 2\mathbf{b}^*$ , hence  $\mathbf{b}^\dagger \in LAT(\nabla)$ ,  $\mathbf{b}^\dagger \neq \mathbf{0}$ .
- From (4.22),

$$LAT(\Delta) = \bigcup_{k \in \mathbb{Z}} LAT(\nabla) + k\mathbf{s}$$

Hence  $\exists \mathbf{b}^\dagger \in LAT(\nabla)$ ,  $\mathbf{b}^\dagger \neq \mathbf{0}$  such that  $\mathbf{0} \in \mathcal{S} \cap (\mathcal{S} + \mathbf{b}^\dagger) \cap LAT(\Delta)$ . Hence  $\mathcal{S} \cap (\mathcal{S} + \mathbf{b}^\dagger) \cap LAT(\Delta) \neq \emptyset$  and by Theorem 4.3,  $LAT(\mathbf{GH}_{LTM})$  is not a  $TS(1)$  sampling lattice on  $LAT(\mathbf{G})$  with respect to  $\mathcal{S}$ , contradicting our assumption. ■

Corollary 4.2 gives the necessary condition for a lattice to be a  $TS(1)$  sampling lattice. Hence, not all lattices in the set  $\mathcal{LE}_n(\mathbf{G})$  are  $TS(1)$  sampling lattices i.e  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \subseteq \mathcal{LE}_n(\mathbf{G})$ . This can be pictorially illustrated in Figure 4.3, for the case of  $N = 2$ .

By Corollary 4.1 and 4.2, we have:

$$\mathcal{LE}_s(\mathbf{G}) \subseteq \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \subseteq \mathcal{LE}_n(\mathbf{G}) \quad (4.33)$$

Note that for the special case of  $N = 1$ , we can arrive at a necessary and sufficient condition to characterize  $TS(1)$  sampling lattices. This is formally stated by the following theorem.

**Theorem 4.4** For  $N = 1$ ,  $LAT(\mathbf{GH}_{LTM})$  with  $\mathbf{G} = \begin{pmatrix} T & 0 \\ 0 & A_1/M_1 \end{pmatrix}$  and  $\mathbf{H}_{LTM} =$

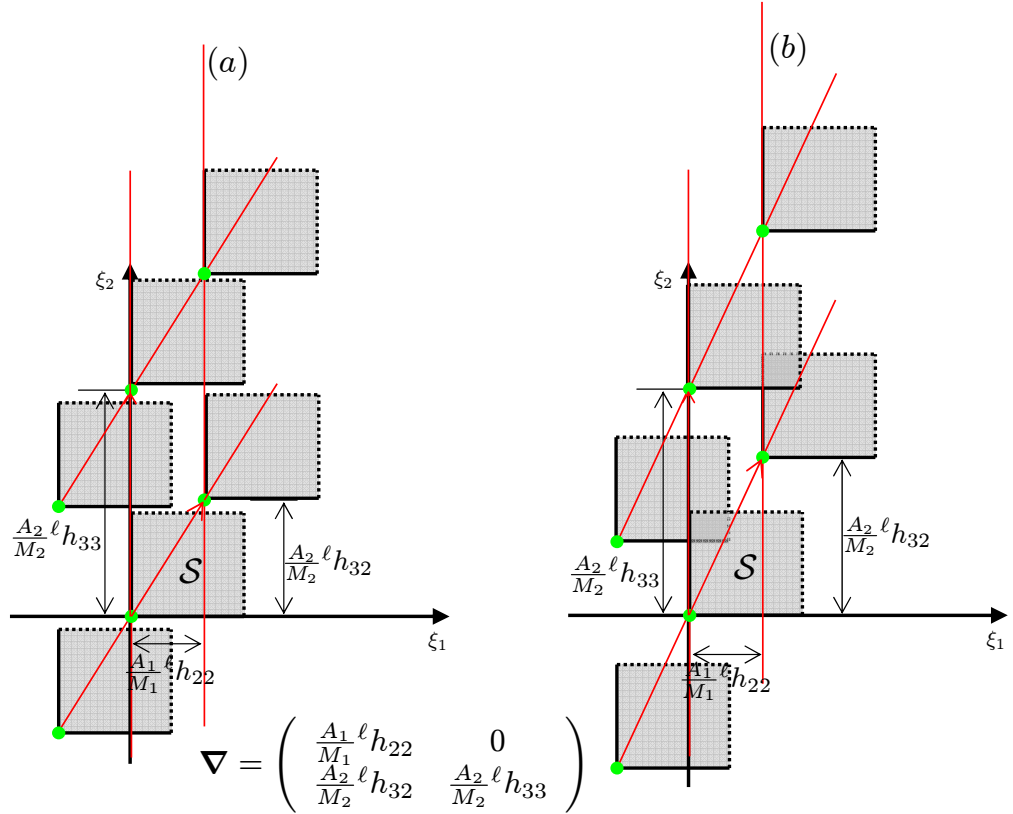


Figure 4.3: The replicas of the spatial support  $\mathcal{S} \subset \mathbb{R}^N$  on  $LAT(\nabla)$ . Note that for  $\ell h_{33} \geq M_2$ , a proper selection of the  $\ell h_{32} \in \mathbb{Z}^+ < \ell h_{33}$  is needed, where in (a) no overlapping of the spatial support replicas occurs whereas in (b) a wrong selection of the  $\ell h_{32}$  results in overlapping of the replicas of the spatial support.

$\begin{pmatrix} 1 & 0 \\ \ell h_{21} & \ell h_{22} \end{pmatrix}$  is a  $TS(1)$  sampling lattice with respect to the spatial domain,  $\mathcal{S} = [0, A_1)$  on grid,  $LAT(\mathbf{G})$  if and only if
 
$$\ell h_{22} \geq M_1 \quad (4.34)$$

**Proof:**

- For  $N = 1$  from (4.24) and (4.31),  $\mathcal{LE}_s(\mathbf{G}) = \mathcal{LE}_n(\mathbf{G})$
- From statement of Corollary 4.1 and 4.2 we arrived at the desired proof. ■

This result form a contribution of this dissertation. This tight condition allows for ease of parameterizing all  $TS(1)$  sampling lattices, hence allowing for fast optimization in finding the optimal  $TS(1)$  sampling lattice for a very important class of temporal-spatial signal (motivated here by the time-varying TRC).

#### 4.3.1 Compactness of the set of non-aliasing $TS(1)$ sampling lattice

From (4.31), the set  $\mathcal{LE}_n(\mathbf{G})$  is not compact. Hence using this set of unique lattices, would make the optimization problem for finding the optimal  $TS(1)$  sampling lattice with respect to  $\mathcal{S} \subset \mathbb{R}^N$  intractable. Fortunately, the problem of finding the optimal  $TS(1)$  sampling lattice requires only those lattices with its polar form that packs  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{1+N}$ .

From (3.10), the compact spectral support,  $\Theta(W, \mathbf{U})$  is assumed to be ellipsoidal. By rescaling the temporal-spatial frequency domain of the sampled signal by the diagonal matrix

$$\mathbf{S}_\Theta = \text{diag}(1/W, 1 \cdot / \mathbf{U}) \quad (4.35)$$

the replication of  $\Theta(W, \mathbf{U})$  on the original temporal-spatial frequency domain of the sampled signal can now be given by the replication of unit balls,  $\mathbf{B} \subset \mathbb{R}^{1+N}$  on the rescaled temporal-spatial frequency domain of the sampled signal. The critical packing lattice that yields the tightest packing of the unit ball,  $\mathbf{B}$  denoted by  $LAT(\mathbf{V}_{crit}^{\mathbf{B}})$  can be found in [54]. Figure 4.4 gives the volume of  $LAT(\mathbf{V}_{crit}^{\mathbf{B}})$  cell,  $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})]$  for different  $(1 + N)$  dimension. Based on these lattices, the critical packing lattice that yields the tightest packing of ellipsoidal body,  $\Theta(W, \mathbf{U})$  is given by  $LAT(\mathbf{V}_{crit}^\Theta)$  where:

$$\mathbf{V}_{crit}^\Theta = \mathbf{S}_\Theta^{-1} \mathbf{V}_{crit}^{\mathbf{B}} \quad (4.36)$$

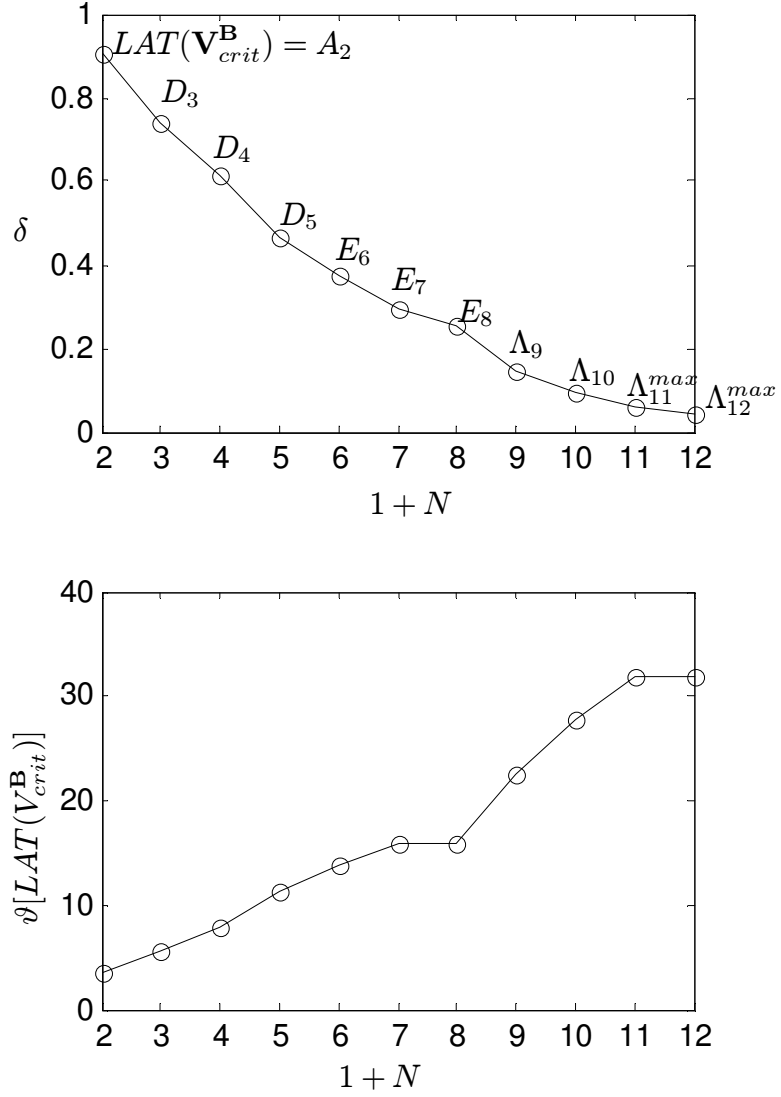


Figure 4.4: The critical packing lattices of the unit ball  $\mathbf{B}$  (top figure) density of the lattice packing,  $\delta = \text{volume of } \mathbf{B} / \vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})]^{\frac{1}{2}}$  for different  $1+N$  dimension; (bottom figure) volume of critical packing cell,  $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})]$  for different  $1+N$  dimension. In particular, for the case of the TRC ( $1+N=2$ ),  $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})] = 2\sqrt{3}$  and for the case of the CRC ( $1+N=4$ ),  $\vartheta[LAT(\mathbf{V}_{crit}^{\mathbf{B}})] = 8$ .



$\mathbf{S}_\Theta^{-1}$  with  $\mathbf{S}_\Theta$  given in (4.35) yields the required scaling matrix to transform the generator matrix from a lattice that packs a unit ball,  $\mathbf{B} \subset \mathbb{R}^{1+N}$  to one that packs an ellipsoidal region,  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{1+N}$ . Since  $LAT(\mathbf{V}_{crit}^\Theta)$  gives the tightest possible packing of the ellipsoidal body,  $\Theta(W, \mathbf{U})$ , the volume of the lattice cell is minimum. By Property 4.4, the volume of the  $LAT(\mathbf{V}_{crit}^\Theta)$  cell is given by:

$$\vartheta[LAT(\mathbf{V}_{crit}^\Theta)] = \left( W \prod_{i=1}^N U_i \right) \vartheta[LAT(\mathbf{V}_{crit}^\mathbf{B})] \quad (4.37)$$

where  $\vartheta[LAT(\mathbf{V}_{crit}^\mathbf{B})]$  is given in Figure 4.4 for different  $(1 + N)$  dimension.

However, typically the sampling lattice described by the critical polar lattice  $LAT(\mathbf{V}_{crit}^\Theta)$ , do not fulfill condition of Theorem 4.3 and therefore is not a  $TS(1)$  sampling lattice. Nonetheless,  $LAT(\mathbf{V}_{crit}^\Theta)$  can be used to deduce a useful sufficient condition for non-aliasing lattices as it specifies the lower bound on packing density. To state this condition, let:

$$\mathcal{P}(\Theta) = \{(\alpha_1(k), T, \mathbf{A}, \mathbf{M}) \mid (\alpha_1(k), T, \mathbf{A}, \mathbf{M}) \text{ gives a non-aliasing sampling}\} \quad (4.38)$$

Hence the non-aliasing lattice condition is given as follows:

**Theorem 4.5** *For a given lattice  $LAT(\mathbf{GH}_{LTM})$  where  $\mathbf{G} = \text{diag}(T, \mathbf{A} \cdot / \mathbf{M})$  and  $\mathbf{H}_{LTM}$  defined by (4.13), if  $LAT(\mathbf{GH}_{LTM})$  belongs to the set  $\mathcal{P}(\Theta) \cap \mathcal{L}(\mathbf{G})$  then*

$$\prod_{i=2}^{1+N} \ell h_{ii} \leq \frac{1}{T \vartheta[LAT(\mathbf{V}_{crit}^\Theta)]} \prod_{i=1}^N \frac{M_i}{A_i} \quad (4.39)$$

**Proof:**

- $LAT(\mathbf{V}_{crit}^\Theta)$  gives the critical polar packing lattice, hence for  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{P}(\Theta) \cap \mathcal{L}(\mathbf{G})$

$$\vartheta[LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)] \geq \vartheta[LAT(\mathbf{V}_{crit}^\Theta)] \quad (4.40)$$

- Since  $\ell h_{11} = 1$  and by lattice Property 4.4, for a given lattice,  $LAT(\mathbf{GH}_{LTM})$ , the volume of its polar form  $\vartheta[LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)]$  where  $\mathbf{G}^* = \mathbf{G}^{-T}$  and  $\mathbf{H}_{LTM} = \mathbf{H}_{LTM}^{-T}$  is given by:

$$\vartheta[LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)] = \prod_{i=1}^N \frac{M_i}{A_i} \bigg/ T \prod_{i=2}^{1+N} \ell h_{ii} \quad (4.41)$$

- Substituting (4.41) into (4.40) and rearranging yields (4.39).

■

By Corollary 4.2 and Theorem 4.5, the set of non-aliasing  $TS(1)$  sampling lattice on grid,  $LAT(\mathbf{G})$  with respect to  $\mathcal{S}$ , is given by:

$$\mathcal{LEP}(\mathbf{G}) = \left\{ LAT(\mathbf{G}\mathbf{H}_{LTM}) \left| \begin{array}{l} \mathbf{H}_{LTM} \text{ by (4.13) takes the values of:} \\ (i) \quad \ell h_{ij} = 0, \text{ for } i < j, \\ (ii) \quad \ell h_{11} = 1 \\ (iii) \quad \ell h_{ii} > 0, \forall i \in \{2, \dots, 1+N\} \\ (iv) \quad \ell h_{ij} \leq 0 \text{ and } |\ell h_{ij}| < \ell h_{ii}, \quad i > j \\ (v) \quad \ell h_{1+N, 1+N} \geq M_N \\ (vi) \quad (\ell h_{22}, \dots, \ell h_{1+N, 1+N}) \text{ fulfills (4.39)} \end{array} \right. \right\} \quad (4.42)$$

Since the set of positive integer combinations  $(\ell h_{22}, \dots, \ell h_{1+N, 1+N})$  fulfilling inequality (4.39) is finite for a given values of  $(T, \mathbf{A}, \mathbf{M})$  parameters (giving us  $LAT(\mathbf{G})$ ), the set  $\mathcal{LEP}(\mathbf{G})$  is compact. This is an important result of this dissertation which suggested that an optimization process that guarantees to terminate can be proposed to find the optimal  $TS(1)$  sampling lattice. Since,

$$\mathcal{LEP}(\mathbf{G}) \supseteq (\mathcal{P}(\Theta) \cap \mathcal{L}(\mathbf{G})) \supset (\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)) \quad (4.43)$$

the idea for the optimization process is to first find if there exist a non-aliasing  $TS(1)$  sampling lattice in the set  $\mathcal{LEP}(\mathbf{G})$  for a given grid,  $LAT(\mathbf{G})$  through an ordered search of the compact set. The existence of the non-aliasing  $TS(1)$  sampling lattice implies that the parameters,  $(T, \mathbf{A}, \mathbf{M})$  that define the grid,  $LAT(\mathbf{G})$  can admit a  $TS(1)$  sampling lattice. We called these grids, *admissible* grids. Then, by optimizing these *admissible* grids such that  $T$  is maximized we obtained the optimal  $TS(1)$  sampling lattice. This will be presented in the next chapter.

#### 4.4 $TS(1)$ Sampling and Lattices

While the lattice theoretic framework provides us a useful mathematical framework for the design of the optimal  $TS(1)$  sampling, not all  $TS(1)$  sampling are lattices. For example the bit-reversed  $TS(1)$  sampling sequence given in Section 3.2.1 (see Figure 3.4) does not admit a lattice structure. However it will be shown that when the temporal-spatial signal<sup>2</sup>,

---

<sup>2</sup>Note that as stated in section 3.2.2 we assume that the  $1+N$  dimensional temporal-spatial signal,  $\omega(t, \mathbf{x})$  is a wide-sense stationary random process.

$\omega(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^N \mapsto \mathbb{R}^N$  is not undersampled temporally by more than a factor of two<sup>3</sup>, if an optimal  $TS(1)$  sampling can be found under this restriction, the optimal sampling lattice will be as good as an unconstrained optimal  $TS(1)$  sampling. This is formally stated by the following theorem. The proof of this theorem is alluded to [38].

**Theorem 4.6** *If  $(\alpha_1(k), T, \mathbf{A}, \mathbf{M})$  is a non-aliasing  $TS(1)$  sampling (not necessarily lattice) with respect to the rectangular spatial domain  $\mathcal{S} = \prod_{i=1}^N [0, A_i]$  with*

$$T \leq 1/(M_t W) \quad (4.44)$$

*then exists a sequence  $\alpha^\diamond(k)$  that is described by a lattice with the same  $T$ .*

**Proof:**

- As stated  $\omega(t, \mathbf{x})$  is the temporally and spatially wide- sense stationary random process. Let  $\omega_s(t, \mathbf{x}) = \omega(kT, \mathbf{X} \cdot * \alpha_1(k))$  be the  $TS(1)$  sampled signal.
- Let,

$$\mathbf{e}(t, \mathbf{x}) := \text{LPF}_\Theta \{ \omega(t, \mathbf{x}) - \omega_s(t, \mathbf{x}) \}$$

be the error between  $\omega(t, \mathbf{x})$  and  $\hat{\omega}(t, \mathbf{x})$  after all frequency components outside of  $\Theta$  have been filtered out by an ideal low pass filter,  $\text{LPF}_\Theta$ .

- The mean-squared error,  $\sigma_e^2$  assuming that  $\omega_s(t, \mathbf{x})$  is also wide-sense stationary and  $\mathbf{e}(t, \mathbf{x})$  and  $\omega(t, \mathbf{x})$  are uncorrelated (see [11]) is given by:

$$\begin{aligned} \sigma_e^2 &= E[|\mathbf{e}(t, \mathbf{x})|^2] \\ &= \int_{\Theta} [\Omega_s(f, \mathbf{u}) - \Omega(f, \mathbf{u})] df d\mathbf{u} \end{aligned} \quad (4.45)$$

- From (3.23) we have:

$$\Omega_s(f, \mathbf{u}) = \sum_{m, \mathbf{p}} |Q_{m\mathbf{p}}|^2 \Omega(f - m/B, \mathbf{u} - \mathbf{p} \cdot / \mathbf{A}) \quad (4.46)$$

---

<sup>3</sup>From temporal Nyquist condition i.e.  $W \leq 1/(2M_t T)$  (given by (3.31)), the temporal-spatial signal,  $\omega(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^N \mapsto \mathbb{R}^N$  is not undersampled temporally by more than a factor of 2 when the inter-sampling time,  $T$  fulfills the following condition:

$$T \leq 1/(M_t W)$$

where  $M_t = \text{prod}(\mathbf{M})$ .

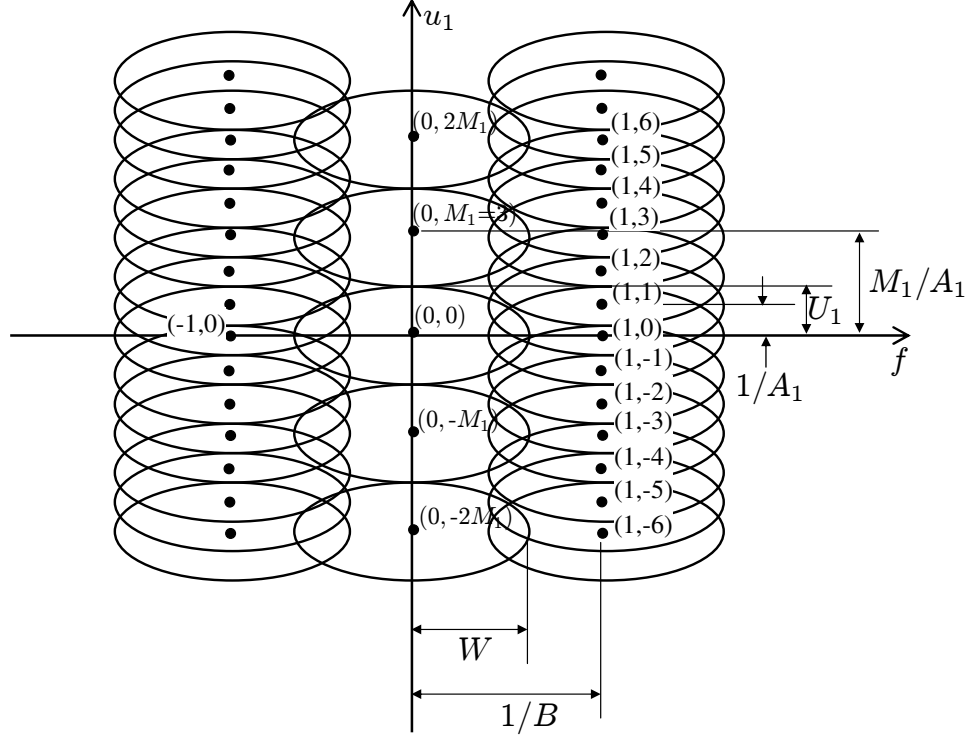


Figure 4.5: The power spectral density of  $TS(1)$  sampled signal of  $1 + N = 2$  temporal-spatial signal given by  $\Omega_s(f, u_1)$  in (4.46). In this figure,  $M_1$  is selected such that it fulfills the spatial Nyquist condition and  $T$  is selected such it go beyond the temporal Nyquist. The black dots indicate the  $(m, p_1)$  position where the spectra replicates. Note that at  $m = 0$  spatial plane, the replicas occurs only at alternate  $M_1$  values due to (4.48).

where  $B = M_t T$ ,  $M_t = \text{prod}(\mathbf{M})$  and the properties of  $Q_{m\mathbf{p}}$  given by (3.25)-(3.28) are:

$$Q_{0\mathbf{0}} = 1 \quad (4.47)$$

$$Q_{0\mathbf{p}} = \delta_{\mathbf{p} \bmod \mathbf{M}} \quad (4.48)$$

$$Q_{m\mathbf{0}} = \delta_{m \bmod M_t} \quad (4.49)$$

$$Q_{m+c*M_t, \mathbf{p}+\mathbf{d}*\mathbf{M}} = Q_{m\mathbf{p}}, \text{ for integer } c \text{ and } \mathbf{d} \quad (4.50)$$

$$Q_{m\mathbf{p}} = Q_{-m, -\mathbf{p}} \quad (4.51)$$

Equation (4.46) shows that the  $TS(1)$  sampled signal spectral,  $\Omega_s(f, \mathbf{u})$  is given by the  $Q_{m\mathbf{p}}$  weighted replicas of the baseband spectra,  $\Omega(f, \mathbf{u})$  at  $(m/B, \mathbf{p} \cdot / \mathbf{A})$  position in the temporal-spatial frequency domain. Figure 4.5 shows an example of the replicas of the baseband spectra,  $\Omega(f, \mathbf{u})$  for  $1 + N = 2$  that pictorially describe equation (4.46).

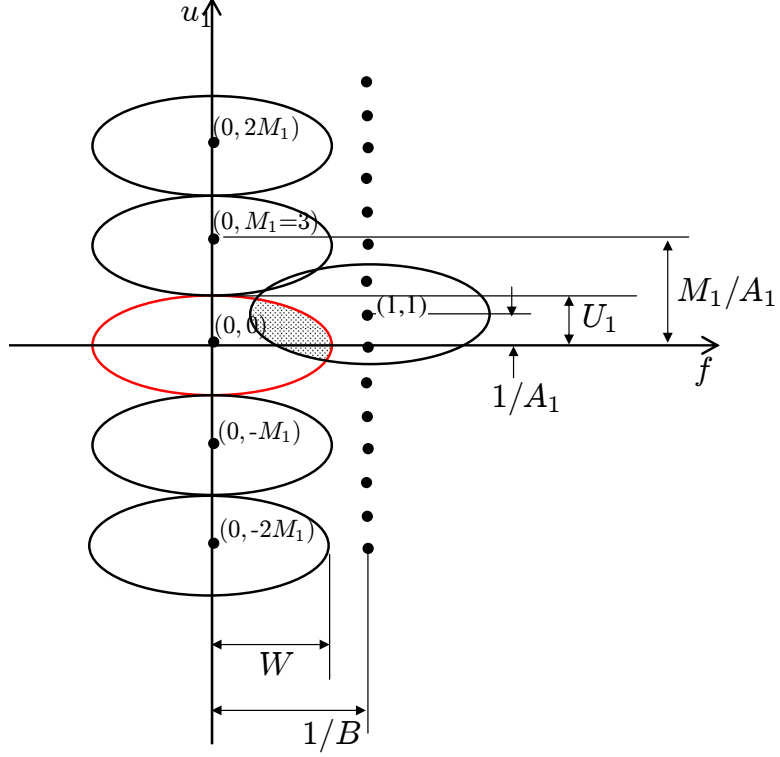


Figure 4.6: Integrating the power spectral density of the replicas at position  $(1, 1)$  on the shaded area gives us  $V_{11}$ .

- Substituting (4.46) into (4.45) we have:

$$\sigma_e^2 = \int_{\Theta} \sum_{(m, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^N / \{\mathbf{0}\}} |Q_{m\mathbf{p}}|^2 \Omega(f - m/B, \mathbf{u} - \mathbf{p} \cdot / \mathbf{A}) + |Q_{00}|^2 \Omega(f, \mathbf{u}) - \Omega(f, \mathbf{u})$$

Applying Property (4.47) we can simplify and rewrite (4.45) to be:

$$\sigma_e^2 = \sum_{(m, \mathbf{p}) \in \mathbb{Z} \times \mathbb{Z}^N / \{\mathbf{0}\}} |Q_{m\mathbf{p}}|^2 V_{m\mathbf{p}} \quad (4.52)$$

where

$$V_{m\mathbf{p}} := \int_{\Theta} \Omega(f - m/B, \mathbf{u} - \mathbf{p} \cdot / \mathbf{A}) df d\mathbf{u} \quad (4.53)$$

$V_{m\mathbf{p}}$  gives the unweighed energy of the spectra replica at position  $(m, \mathbf{p})$  that overlaps the baseband region given by  $\Theta$ . An example describing (4.53) is pictorially shown in Figure 4.6.

- By the spatial Nyquist condition given in (3.30),  $(\boldsymbol{\alpha}_1(k), T, \mathbf{A}, \mathbf{M}) \in \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$

only if we pick

$$\mathbf{M} \geq 2\mathbf{A} \cdot * \mathbf{U} \quad (4.54)$$

Note that with this condition, spectra replicas for all  $|\mathbf{p}| \geq \mathbf{M}$  at all  $m$  will not overlap with the baseband. This implies that:

$$V_{m\mathbf{p}} = 0 \text{ for } |\mathbf{p}| \geq \mathbf{M} \quad (4.55)$$

Also from condition (4.44),

$$T \leq 1/(M_t W) \quad (4.56)$$

Note that with this condition, spectra replicas at  $|m| > 1$  for all  $\mathbf{p}$  will not overlap with the baseband. This implies that:

$$V_{m\mathbf{p}} = 0 \text{ for } |m| > 1 \quad (4.57)$$

- By (4.55) and (4.57), (4.52) can be further simplified to be:

$$\begin{aligned} \sigma_e^2 &= \sum_{m=-1}^1 \sum_{\mathbf{p}=-\mathbf{M}+1}^{\mathbf{M}-1} |Q_{m\mathbf{p}}|^2 V_{m\mathbf{p}} - |Q_{00}|^2 V_{00} \\ &= 2 \sum_{\mathbf{p}=0}^{\mathbf{M}-1} |Q_{1\mathbf{p}}|^2 W_{\mathbf{p}} \end{aligned} \quad (4.58)$$

where

$$W_{\mathbf{p}} = \sum_{\mathbf{b}=0}^1 V_{1,\mathbf{p}-\mathbf{b}\mathbf{M}} \quad (4.59)$$

with the sum given over the  $2^N$   $N$ -tuples  $\mathbf{b}$  containing  $\mathbf{0}$ 's and  $\mathbf{1}$ 's.

- The  $Q_{m\mathbf{p}}$  in (3.24) was derived by specifying the sampling sequence in terms of the spatial coordinate  $\alpha_1(k)$  where the  $k$ -th sample in the sequence is taken for  $k = 0, \dots, M_t - 1$ . We may alternatively define the sequence in terms of the position  $0 \leq \gamma(\mathbf{a}) \leq M_t - 1$  of the spatial coordinates  $\mathbf{a}$  in the time sequence for  $\mathbf{a} = \mathbf{0}, \dots, \mathbf{M} - \mathbf{1}$ . In this case, (3.24) becomes:

$$Q_{m\mathbf{p}} = \frac{1}{M_t} \sum_{\mathbf{a}=0}^{\mathbf{M}-1} \exp^{-i2\pi[m\gamma(\mathbf{a})/M_t + \mathbf{p} \cdot * (\mathbf{a}/\mathbf{M})]}$$

For each value of  $m$ , this is an  $N$ -dimensional DFT with respect to  $\mathbf{a}$  of the signal:

$$q_{m\mathbf{a}} = \exp^{-i2\pi m\gamma(\mathbf{a})/M_t} \quad (4.60)$$

By Parseval's relation for the DFT:

$$\sum_{\mathbf{p}=0}^{\mathbf{M}-1} |Q_{m\mathbf{p}}|^2 = 1, \text{ for each } m \quad (4.61)$$

- Now from (4.58) and (4.61), taking  $m = 1$ , the non-aliasing  $TS(1)$  sequence is one that will distribute 1 unit of energy over the coefficients  $Q_{1\mathbf{p}}$  in a way that their weighted sum given by (4.58) equals to 0 (i.e.  $\sigma_{\mathbf{e}}^2 = 0$ ). By the discussion given in section 3.2.2, we can easily achieved this by fulfilling the temporal Nyquist condition i.e.  $T \leq 1/(2M_tW)$  since  $W_{\mathbf{p}} = 0$  for each  $\mathbf{p}$  in this case.
- However, the temporal Nyquist rate may be less than  $2W$  if  $Q_{m\mathbf{p}} = 0$  for values of  $(m, \mathbf{p})$  other than those fixed by  $Q_{m\mathbf{0}} = \delta_{\mathbf{p} \cdot \text{mod } \mathbf{M}}$  (given by (4.49)). This can be achieved by selecting the “right”  $TS(1)$  sampling sequence,  $\alpha_1(k)$ . As the inter-sampling time increases beyond the Nyquist rate i.e.  $T \geq 1/(2M_tW)$ , we always assigns all the energy to the  $\mathbf{p}$  spatial position where  $W_{\mathbf{p}} = 0$  i.e. we choose:

$$Q_{1\mathbf{p}} = \begin{cases} 1 & \text{if } p = p^\diamond \\ 0 & \text{otherwise} \end{cases} \quad (4.62)$$

with  $W_{\mathbf{p}^\diamond} = 0$ . From (4.50),

$$Q_{1\mathbf{p}} = Q_{1, \mathbf{p} + \mathbf{d} \cdot \mathbf{M}} \quad (4.63)$$

Equations (4.62) and (4.63) can be written compactly as:

$$Q_{1\mathbf{p}} = \delta_{(\mathbf{p} - \mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}} \quad (4.64)$$

- To construct  $Q_{m\mathbf{p}}$  from  $Q_{1\mathbf{p}}$  consider (4.60)

$$q_{m\mathbf{a}} = [q_{1\mathbf{a}}]^m \quad (4.65)$$

Taking the DFT on (4.65) and by the convolution theorem we have,

$$\begin{aligned} \text{DFT}[q_{m\mathbf{a}}] &= \text{DFT}[q_{1\mathbf{a}}] * \text{DFT}[q_{1\mathbf{a}}] * \dots * \text{DFT}[q_{1\mathbf{a}}] \\ Q_{m\mathbf{p}} &= [Q_{1\mathbf{p}}]_{*m} \end{aligned} \quad (4.66)$$

where  $[\cdot]_{*m}$  denotes the  $m$ -fold self convolution. Now, substituting (4.64) into (4.66) we have:

$$Q_{m\mathbf{p}} = [\delta_{(\mathbf{p} - \mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}}]_{*m} \quad (4.67)$$

Let  $\underline{\delta}[\mathbf{p}] := \delta_{(\mathbf{p}-\mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}}$ . Note  $\underline{\delta}[\mathbf{p} + \mathbf{r} \cdot * \mathbf{M}] = \underline{\delta}[\mathbf{p}]$  for  $\mathbf{r} \in \mathbb{Z}^N$  because  $(\mathbf{p} + \mathbf{r} \cdot * \mathbf{M} - \mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M} = (\mathbf{p} - \mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}$ .  $\underline{\delta}[\mathbf{p}]$  is a periodic sequence. By circular convolution:

$$\begin{aligned} \underline{\delta}[\mathbf{p}] * \underline{\delta}[\mathbf{p}] &= \sum_{\bar{\mathbf{p}}=0}^{\mathbf{M}-1} \underline{\delta}[\bar{\mathbf{p}}] \underline{\delta}[\mathbf{p} - \bar{\mathbf{p}}] \\ &= \underline{\delta}[\mathbf{p} - \mathbf{p}^\diamond] \\ &= \delta_{(\mathbf{p}-2\mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}} \end{aligned}$$

Hence, by induction from (4.67) we have:

$$Q_{m\mathbf{p}} = \delta_{(\mathbf{p}-m\mathbf{p}^\diamond) \cdot \text{mod } \mathbf{M}} \quad (4.68)$$

- Equation (4.68) implies that  $Q_{m\mathbf{p}}$  is described by a lattice. Since the  $Q_{m\mathbf{p}}$  dictates the frequency content replication of the  $TS(1)$  sampled signal this means that the frequency content replication is described by a lattice. By Property 4.1 of lattices,  $\alpha^\diamond(k)$  is a lattice.

Note that for this theorem to be valid, the temporal-spatial signal,  $\omega(t, \mathbf{x}) : \mathbb{R} \times \mathbb{R}^N \mapsto \mathbb{R}^N$  cannot be undersampled temporally by more than a factor of 2 i.e.  $T \leq 1/(M_t W)$ . This is a restriction since we cannot make a statement on the optimality of the designed  $TS(1)$  sampling lattice if the optimized inter-sampling time is more than  $1/(M_t W)$ . ■

However, it turns out that for the special case of the time-varying TRC i.e.  $1 + N = 2$  dimensional temporal-spatial signal with assumed spectral support given by the ellipsoidal body,  $\Theta(W, U_1)$ , the designed  $TS(1)$  sampling lattice is indeed optimal for all sampling forms (lattice and non-lattice). This is because for this case a non-aliasing  $TS(1)$  sampling (lattice or non-lattice) can be shown to have an inter-sampling time  $T \leq 1/(M_1 W)$ . Therefore Theorem 4.6 guarantees to be valid in this case. This is stated by the following corollary.

**Corollary 4.3** *For  $1 + N = 2$  dimensional temporal-spatial signal with spectral support given by the ellipsoidal body,  $\Theta(W, U_1)$  if  $(\alpha_1(k), T, A_1, M_1)$  is a non-aliasing  $TS(1)$  sampling (not necessarily lattice) with respect to the rectangular spatial domain  $\mathcal{S} = [0, A_1]$  then exists a sequence  $\alpha^\diamond(k)$  that is described by a lattice with the same  $T$ .*

**Proof:**



- From Theorem 4.5, for  $1 + N = 2$ ,

$$T \leq \frac{1}{2\sqrt{3}WU_1A_1} \quad (4.69)$$

- Since for non-aliasing  $TS(1)$  sampling the spatial Nyquist condition given by (3.30) must be fulfilled, we have:

$$\frac{M_1}{A_1} \geq 2U_1 \quad (4.70)$$

- Together with (4.69), we have:

$$T \leq \frac{1}{\sqrt{3}M_1W} < \frac{1}{M_1W} \quad (4.71)$$

- By Theorem 4.6, we arrived at the desired result.

■

## 4.5 Summary

In this chapter, a lattice based framework is given for parameterizing  $TS(1)$  sampling lattices on a grid,  $LAT(\mathbf{G})$  with respect to the spatial support,  $\mathcal{S}$ . It is shown here that the set of all non-aliasing  $TS(1)$  sampling lattice on grid,  $LAT(\mathbf{G})$  with respect to  $\mathcal{S}$  is compact. Hence subject to the assumption that the maximum resolution of the spatial grid is upper bounded, an optimization process that guarantee to terminate can be proposed to find the optimal  $TS(1)$  sampling lattice. This optimization process will be presented in the next chapter. While the lattice theoretic framework provides us a useful mathematical framework for the design of the optimal  $TS(1)$  sampling, not all  $TS(1)$  sampling are lattices. It is shown here that when the temporal-spatial signal is not undersampled temporally by more than a factor of 2, if an optimal  $TS(1)$  sampling can be found under this restriction, the optimal lattice sampling will be as good as any unconstrained optimal  $TS(1)$  sampling.

## Chapter 5

# Design of Time-Sequential Sampling

Time-sequential sampling is described in the lattice framework in the previous chapter. The set of non-aliasing  $TS(1)$  sampling lattice on a given grid,  $LAT(\mathbf{G})$  with respect to the spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$  given by  $\mathcal{LEP}(\mathbf{G})$  is compact. In addition the parameters that defined the grid,  $(T, \mathbf{A}, \mathbf{M})$  (see equation (4.10)) are all bounded. This important result is given in section 5.3 of this chapter. Therefore an optimization process that guarantees to terminate can be proposed to find the  $TS(1)$  sampling lattice with the highest inter-sampling time,  $T$  that avoids aliasing of the assumed compact ellipsoidal spectral support. This is called the optimal  $TS(1)$  sampling lattice. From the optimal  $TS(1)$  sampling lattice, we can then construct the  $TS(n)$  sampling lattice for  $n > 1$ .

Section 5.1, gives an optimization approach to find the optimal  $TS(1)$  sampling lattice. Section 5.2 discussed the algorithm to find a non-aliasing  $TS(1)$  sampling lattice for a given grid,  $LAT(\mathbf{G})$ . Then the optimization algorithm to find the optimal  $TS(1)$  sampling lattice is given in section 5.3. To reduce the computational complexity of the optimization algorithm for finding the optimal  $TS(1)$  sampling lattice, lattice restriction conditions are proposed in section 5.4. The extension to the design of the sub-optimal  $TS(n), n > 1$  sampling lattice is given in section 5.5. In section 5.6, case studies for the  $TS(n)$  sampling design is given for both the tone reproduction curve, TRC and the color reproduction characteristics function, CRC. The effectiveness and performance gain with this design compared to other consistent sampling approaches are given in this section.

## 5.1 Optimization of $TS(1)$ Sampling Lattices

In this section, we proposed an optimization approach to find the  $TS(1)$  sampling lattice with the highest inter-sampling time,  $T$  that avoids aliasing of the assumed compact ellipsoidal spectral support,  $\Theta$ . From (4.17),  $\mathcal{L}(\mathbf{G})$  defines the set of all unique lattices on the grid,  $LAT(\mathbf{G})$ . Let

$$\mathcal{L}_{grid} = \left\{ LAT(\mathbf{G}) \left| \mathbf{G} = \begin{pmatrix} T & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & A_N/M_N \end{pmatrix}, \forall T \in \mathbb{R}^+, \forall \mathbf{M} \in \mathbb{Z}^{+N} \text{ and } \mathbf{A} \in \mathbb{R}^{+N} \right. \right\} \quad (5.1)$$

$\mathcal{L}_{grid}$  denotes the set of all possible grids. Then, we can define the set of all lattices as:

$$\mathcal{L} = \bigcup_{LAT(\mathbf{G}) \in \mathcal{L}_{grid}} \mathcal{L}(\mathbf{G}) \quad (5.2)$$

Hence, from the definition of the  $TS(1)$  sampling set and the non-aliasing sampling set given by (4.18) and (4.38) respectively, the set of all non-aliasing  $TS(1)$  sampling lattices with respect to the spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$  can now be given by:

$$\mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta) \quad (5.3)$$

Let  $\mathcal{T}_{LAT}(1, \mathcal{S}, \Theta)$  be the set of inter-sampling times of all the non-aliasing  $TS(1)$  sampling lattices i.e.:

$$\mathcal{T}_{LAT}(1, \mathcal{S}, \Theta) = \{T | (\alpha_1(k), T, \mathbf{A}, \mathbf{M}) \in \mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)\} \quad (5.4)$$

The optimal  $TS(1)$  sampling lattice is defined as follows.

**Definition 5.1 (optimal  $TS(1)$  sampling lattice)** *A  $TS(1)$  sampling lattice given by  $(\alpha_1^*(k), T^*, \mathbf{A}, \mathbf{M}^*)$  with  $\mathbf{A}$  known is optimal, if*

$$(i) (\alpha_1^*(k), T^*, \mathbf{A}, \mathbf{M}^*) \in \mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$$

$$(ii) T^* \geq T, \forall T \in \mathcal{T}_{LAT}(1, \mathcal{S}, \Theta)$$

By Definition 5.1, the optimization problem to find the optimal  $TS(1)$  sampling lattice can now be stated as follows:

**Problem 5.1** *Find  $LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)$  that belongs to the set  $\mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  where  $\mathbf{G}^* = \text{diag}(T^*, \mathbf{A} \cdot / \mathbf{M}^*)$  such that  $T^* \geq T, \forall T \in \mathcal{T}_{LAT}(1, \mathcal{S}, \Theta)$ .*

The optimization approach to solve Problem 5.1 makes use of the fact that it is possible to decompose the set of all non-aliasing  $TS(1)$  sampling lattices,  $\mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  into a collection of non-aliasing  $TS(1)$  sampling lattices on grid,  $LAT(\mathbf{G})$  that is:

$$\mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta) = \bigcup_{LAT(\mathbf{G}) \in \mathcal{L}_{grid}} \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta) \quad (5.5)$$

The idea then is to look at each grid individually and find a non-aliasing  $TS(1)$  sampling lattice. Grids that yield a non-aliasing  $TS(1)$  sampling lattice are admissible grids. Note that these admissible grids depends on the inter-sampling time,  $T$  and the resolution of the spatial domain,  $\mathbf{X} = \mathbf{A} \cdot / \mathbf{M}$  (see Figure 3.3). By optimizing these admissible grids such that the inter-sampling time,  $T$  is maximized, by statement of Problem 5.1, we obtained the optimal  $TS(1)$  sampling lattice. The algorithm to perform this optimization is given in section 5.3.

## 5.2 Non-aliasing $TS(1)$ Sampling Lattice for a Given Grid, $LAT(\mathbf{G})$

The set of non-aliasing  $TS(1)$  sampling lattices given by  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  is defined by the known compact spectral support,  $\Theta(W, \mathbf{U})$ ; the known rectangular spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$ ; and the grid,  $LAT(\mathbf{G})$ . From Chapter 4, the set  $\mathcal{LEP}(\mathbf{G})$  given by (4.42) gives the superset of the non-aliasing  $TS(1)$  sampling lattice i.e. from (4.43):

$$\mathcal{LEP}(\mathbf{G}) \supset (\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)) \quad (5.6)$$

Notice that a non-empty  $\mathcal{LEP}(\mathbf{G})$  does not imply that  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  is non-empty. In fact, there can be lattices in  $\mathcal{LEP}(\mathbf{G})$  which are neither  $TS(1)$  sampling lattice nor results in packing of the ellipsoidal spectral support.

However, since  $\mathcal{LEP}(\mathbf{G})$  is compact, an ordered search can be performed on lattices in this set to find a non-aliasing  $TS(1)$  sampling lattice. In order to perform this search, it is necessary to have algorithms to determine if a given lattice,  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing  $TS(1)$  sampling lattice. To achieve this we first check if  $LAT(\mathbf{GH}_{LTM})$  belongs to  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ , then check if  $LAT(\mathbf{GH}_{LTM})$  belongs to the set  $\mathcal{P}(\Theta)$ .

### 5.2.1 Check if a given lattice, $LAT(\mathbf{GH}_{LTM})$ is a $TS(1)$ sampling lattice

To check if  $LAT(\mathbf{GH}_{LTM})$  where  $\mathbf{GH}_{LTM} = \begin{pmatrix} T & 0 \\ \mathbf{s} & \nabla \end{pmatrix} \in \mathbb{R}^{(1+N) \times (1+N)}$  is a  $TS(1)$  sampling lattice, we make use of Theorem 4.3. First, define the following integer set:

$$\mathcal{V}_N = \{\ell_{\mathbf{z}_s} \mid \mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta) \neq \emptyset, \mathbf{b} = \nabla^\ell \mathbf{z}_s\} \subset \mathbb{Z}^N \quad (5.7)$$

where  $LAT(\Delta) = \bigcup_{k \in \mathbb{Z}} LAT(\nabla) + ks$ . By Theorem 4.3, for a given spatial support,  $\mathcal{S} \subset \mathbb{R}^N$ , the lattice  $LAT(\mathbf{GH}_{LTM})$  belongs to the set  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$  (i.e. it is a  $TS(1)$  sampling lattice) iff  $\mathcal{V}_N = \{\mathbf{0}\}$ . Hence, to check if a given lattice,  $LAT(\mathbf{GH}_{LTM})$  yields a  $TS(1)$  sampling lattice requires construction of the set  $\mathcal{V}_N \subset \mathbb{Z}^N$  as defined in (5.7). To construct this integer set consider the following definition of integer sets:

$${}^\alpha \mathcal{V}_N = \{\ell_{\mathbf{z}_s} \mid \mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \neq \emptyset, \mathbf{b} = \nabla^\ell \mathbf{z}_s\} \subset \mathbb{Z}^N \quad (5.8)$$

and,

$${}^\beta \mathcal{V}_N(\ell_{\mathbf{z}_s}) = \{u_{\mathbf{z}_s} \mid \Delta^u \mathbf{z}_s \in \mathcal{S} \cap (\mathcal{S} + \nabla^\ell \mathbf{z}_s)\} \subset \mathbb{Z}^N \quad (5.9)$$

Note that for  $\ell_{\mathbf{z}_s} \in {}^\alpha \mathcal{V}_N$ ,  ${}^\beta \mathcal{V}_N(\ell_{\mathbf{z}_s})$  non-empty means that exists a  $\ell_{\mathbf{z}_s} \in \mathbb{Z}^N$  which resulted in  $\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta) \neq \emptyset$  for  $\mathbf{b} = \nabla^\ell \mathbf{z}_s$ . Hence, (5.8) and (5.9) defines  $\mathcal{V}_N \subset \mathbb{Z}^N$  given by (5.7) i.e.:

$$\mathcal{V}_N = \left\{ \ell_{\mathbf{z}_s} \mid \ell_{\mathbf{z}_s} \in {}^\alpha \mathcal{V}_N \text{ with } {}^\beta \mathcal{V}_N(\ell_{\mathbf{z}_s}) \neq \emptyset \right\} \subset \mathbb{Z}^N \quad (5.10)$$

To construct  $\mathcal{V}_N \subset \mathbb{Z}^N$  we need to first find the integer set  ${}^\alpha \mathcal{V}_N \subset \mathbb{Z}^N$  and for every  $\ell_{\mathbf{z}_s} \in {}^\alpha \mathcal{V}_N$  find the integer set  ${}^\beta \mathcal{V}_N(\ell_{\mathbf{z}_s})$  as follows:

#### 5.2.1.1 Find integer space ${}^\alpha \mathcal{V}_N \subset \mathbb{Z}^N$

- Since  $\mathcal{S} = \prod_{i=1}^N [0, A_i)$ , for  $\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \neq \emptyset$  the vector components of  $\mathbf{b} \in LAT(\nabla)$  are given by:

$$-A_i < b_i < A_i, \forall i \in \{1, \dots, N\} \quad (5.11)$$

Since  $\mathbf{b} \in LAT(\nabla)$ , from (4.15) we have,

$$\underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}}_{\mathbf{b}} = \underbrace{\begin{pmatrix} \frac{A_1}{M_1} \ell h_{22} & 0 & \dots & 0 \\ \frac{A_2}{M_2} \ell h_{32} & \frac{A_2}{M_2} \ell h_{33} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{A_N}{M_N} \ell h_{1+N,2} & \frac{A_N}{M_N} \ell h_{1+N,3} & \dots & \frac{A_N}{M_N} \ell h_{1+N,1+N} \end{pmatrix}}_{\nabla} \underbrace{\begin{pmatrix} \ell z_2 \\ \ell z_3 \\ \vdots \\ \ell z_{1+N} \end{pmatrix}}_{\ell_{\mathbf{z}_s}} \quad (5.12)$$

Note that the diagonal elements of  $\nabla$  are always greater than 0 i.e.  $\nabla_{ii} > 0$  since  ${}^\ell h_{i+1,i+1} > 0$  (by HNF theorem),  $M_i > 0$  and  $A_i > 0$  for all  $i$ . Thus, in the development to follow, division of a finite value by  $\nabla_{ii}$  would not cause a reversal of its sign and its value will be finite.

The idea in finding the integer set  ${}^\alpha \mathcal{V}_N \subset \mathbb{Z}^N$  defined by (5.8) is to make use of the lower triangular form of  $\nabla \in \mathbb{R}^{N \times N}$  and successively find the bound for elements of integer vector  ${}^\ell z_s \in \mathbb{Z}^N$  such that (5.11) is violated.

- Starting with  $i = 1$ ,  $b_1 = \nabla_{11} {}^\ell z_2$ . To have  $b_1 \in \mathbb{R}$  in the range of  $-A_1 < b_1 < A_1$  means that  ${}^\ell z_2 \in \mathbb{Z}$  is in the range of:

$$\left\lfloor -\frac{A_1}{\nabla_{11}} \right\rfloor + 1 \leq {}^\ell z_2 \leq \left\lceil \frac{A_1}{\nabla_{11}} \right\rceil - 1$$

Let,

$${}^\alpha \mathcal{V}_1 = \left\{ {}^\ell z_2 \mid \left\lfloor -\frac{A_1}{\nabla_{11}} \right\rfloor + 1 \leq {}^\ell z_2 \leq \left\lceil \frac{A_1}{\nabla_{11}} \right\rceil - 1 \right\} \subset \mathbb{Z} \quad (5.13)$$

Note that since  $A_1 > 0$  and  $\nabla_{11} > 0$ ,  ${}^\alpha \mathcal{V}_1 \subset \mathbb{Z}$  is non-empty with at least 0 as its element.

- For  $i = 2$ ,  $b_2 = \nabla_{21} {}^\ell z_2 + \nabla_{22} {}^\ell z_3$ . To have  $b_2 \in \mathbb{R}$  in the range of  $-A_2 < b_2 < A_2$  means that  ${}^\ell z_3 \in \mathbb{Z}$  is in the range of

$$\left\lfloor \frac{-A_2 - \nabla_{21} {}^\ell z_2}{\nabla_{22}} \right\rfloor + 1 \leq {}^\ell z_3 \leq \left\lceil \frac{-A_2 - \nabla_{21} {}^\ell z_2}{\nabla_{22}} \right\rceil - 1 \quad (5.14)$$

Let

$${}^\alpha \mathcal{V}_2 = \left\{ ({}^\ell z_2, {}^\ell z_3) \mid \begin{array}{l} \forall {}^\ell z_2 \in {}^\alpha \mathcal{V}_1, \\ {}^\ell z_3 \text{ fulfills inequality (5.14)} \end{array} \right\} \subset \mathbb{Z} \times \mathbb{Z}$$

- In general, for  $i > 1$ , to fulfill  $-A_i < b_i < A_i$ , we need  ${}^\ell z_{i+1} \in \mathbb{Z}$  to be within the range of

$$\left\lfloor \frac{-A_i - \sum_{j=1}^{i-1} \nabla_{ij} {}^\ell z_{j+1}}{\nabla_{ii}} \right\rfloor + 1 \leq {}^\ell z_{i+1} \leq \left\lceil \frac{A_i - \sum_{j=1}^{i-1} \nabla_{ij} {}^\ell z_{j+1}}{\nabla_{ii}} \right\rceil - 1 \quad (5.15)$$

Let,

$${}^\alpha \mathcal{V}_i = \left\{ ({}^\ell z_2, {}^\ell z_3, \dots, {}^\ell z_{i+1}) \mid \begin{array}{l} \forall ({}^\ell z_2, \dots, {}^\ell z_i) \in {}^\alpha \mathcal{V}_{i-1}, \\ {}^\ell z_{i+1} \text{ fulfills (5.15)} \end{array} \right\} \quad (5.16)$$

For  $i = 2$  to  $i = N$ , recursively compute  ${}^\alpha \mathcal{V}_i$  from (5.16). At  $i = N$ , we have

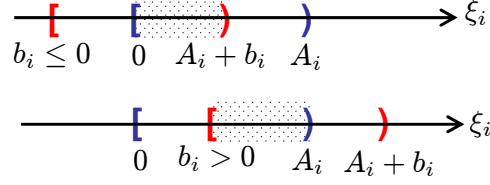


Figure 5.1: Two forms of overlapped region at each  $i$ -th spatial dimension

$${}^{\alpha}\mathcal{V}_N \subset \mathbb{Z}^N.$$

Note that in this development:

1. Since  ${}^{\alpha}\mathcal{V}_1 \subset \mathbb{Z}$  is non-empty with at least 0 as its element,  ${}^{\alpha}\mathcal{V}_N \subset \mathbb{Z}^N$  is non-empty and contain at least the  ${}^{\ell}\mathbf{z}_s = \mathbf{0}$  as its element.
2. Also, since the bound for  ${}^{\ell}z_{i+1}$  for each  $i = 2, \dots, N$  from (5.15) is finite ( $\nabla_{ii} > 0$ ), hence the set  ${}^{\alpha}\mathcal{V}_N$  contains finite number of elements.

5.2.1.2 Find integer space  ${}^{\beta}\mathcal{V}_N({}^{\ell}\mathbf{z}_s)$  for every  ${}^{\ell}\mathbf{z}_s \in {}^{\alpha}\mathcal{V}_N$

The integer space  ${}^{\beta}\mathcal{V}_N({}^{\ell}\mathbf{z}_s)$  defined by (5.9) gives the set of all integer vector,  ${}^u\mathbf{z}_s$  such that the point  $\Delta^u\mathbf{z}_s$  is in the overlapped region  $\mathcal{S} \cap (\mathcal{S} + \nabla^{\ell}\mathbf{z}_s)$ ,  ${}^{\ell}\mathbf{z}_s \in {}^{\alpha}\mathcal{V}_N$

Since  $\mathcal{S} \subset \mathbb{R}^N$  is a rectangular domain, the overlapped spatial region,

$$\mathcal{O}({}^{\ell}\mathbf{z}_s) := \mathcal{S} \cap (\mathcal{S} + \nabla^{\ell}\mathbf{z}_s) \subseteq \mathcal{S} \quad (5.17)$$

can be described by looking at the overlapped region at each  $i$ -th spatial dimension. The  $i$ -dimension overlapped region is given in two forms as shown in Figure 5.1.

For  $\mathbf{b} := [b_1, \dots, b_N]^T = \nabla^{\ell}\mathbf{z}_s$ ,  ${}^{\ell}\mathbf{z}_s \in {}^{\alpha}\mathcal{V}_N$ , the overlapped region,  $\mathcal{O}({}^{\ell}\mathbf{z}_s) \subset \mathbb{R}^N$  is given by:

$$\mathcal{O}({}^{\ell}\mathbf{z}_s) = \prod_{i=1}^N [{}^{\ell}s_i, {}^us_i) \quad (5.18)$$

where,

$${}^{\ell}s_i = \begin{cases} 0 & \text{if } b_i \leq 0 \\ b_i & \text{if } b_i > 0 \end{cases}, \quad {}^us_i = \begin{cases} A_i + b_i & \text{if } b_i \leq 0 \\ A_i & \text{if } b_i > 0 \end{cases} \quad (5.19)$$

For example, for the case of  $N = 2$  spatial dimension, the overlapped region for  $\mathcal{O}({}^{\ell}\mathbf{z}_s) \subset \mathbb{R} \times \mathbb{R}$  is as shown in Figure 5.2. Now, to find the integer vector  ${}^u\mathbf{z}_s$  such that the point

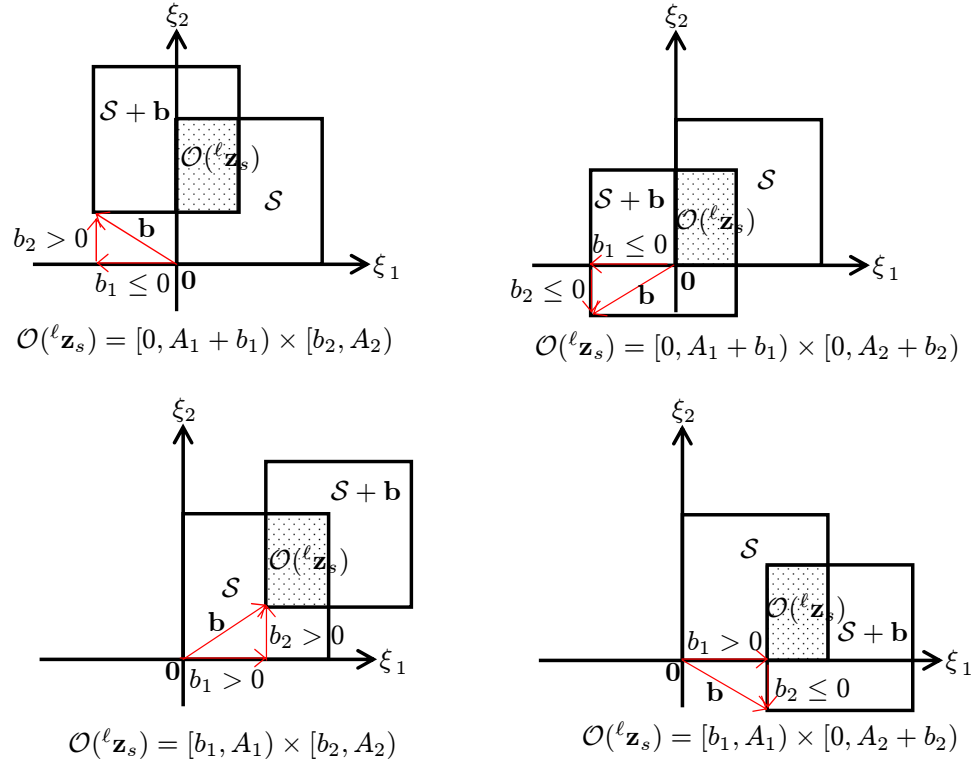


Figure 5.2: Possible overlapped region configuration of  $\mathcal{O}(\ell_{\mathbf{z}_s})$  for  $N=2$



$\Delta^{\ell \mathbf{z}_s}$  is in  $\mathcal{O}(\ell \mathbf{z}_s)$ , the following algorithm is used:

- Let  $\zeta_s = \Delta^u \mathbf{z}_s \in \mathbb{R}^N$  with  ${}^u \mathbf{z}_s := [{}^u z_2, \dots, {}^u z_{1+N}]^T \in \mathbb{Z}^N$ . For  $\mathcal{O}(\ell \mathbf{z}_s) \cap \zeta_s \neq \emptyset$ , the vector components of  $\zeta_s \in \mathbb{R}^N$  are given by:

$${}^{\ell} s_i \leq \zeta_{s,i} < {}^u s_i, \forall i \in \{1, \dots, N\} \quad (5.20)$$

The idea of finding the integer set  ${}^{\beta} \mathcal{V}_N(\ell \mathbf{z}_s) \subset \mathbb{Z}^N$  is to make use of the upper triangular form of  $\Delta \in \mathbb{R}^{N \times N}$  and successively find the bound for elements of integer vectors  ${}^u \mathbf{z}_s$  such that (5.20) is violated.

Since  $\zeta_s = \Delta^u \mathbf{z}_s \in \mathbb{R}^N$  with  ${}^u \mathbf{z}_s := [{}^u z_2, \dots, {}^u z_{1+N}]^T \in \mathbb{Z}^N$ , we have:

$$\zeta_{s,i} = \sum_{j=i}^N \Delta_{ij} {}^u z_{j+1} \in \mathbb{R}, \quad i = 1, \dots, N \quad (5.21)$$

Note that from section 4.3  $\Delta_{ii} > 0$  for all  $i$ -th spatial dimension. Thus, in the development to follow, division of a finite value by  $\Delta_{ii}$  would not cause a reversal of its sign and its value will be finite.

- Starting with  $i = N$ ,  $\zeta_{s,N} = \Delta_{NN} {}^u z_{1+N} \in \mathbb{R}$ . Hence, to have  $\zeta_{s,N}$  in the range of  ${}^{\ell} s_N \leq \zeta_{s,N} < {}^u s_N$  means that  ${}^u z_{1+N} \in \mathbb{Z}$  is in the range of:

$$\left\lceil \frac{{}^{\ell} s_N}{\Delta_{NN}} \right\rceil \leq {}^u z_{1+N} \leq \left\lfloor \frac{{}^u s_N}{\Delta_{NN}} \right\rfloor - 1 \quad (5.22)$$

Let,

$${}^{\beta} \mathcal{V}_1(\ell \mathbf{z}_s) = \left\{ {}^u z_{1+N} \left| \left\lceil \frac{{}^{\ell} s_N}{\Delta_{NN}} \right\rceil \leq {}^u z_{1+N} \leq \left\lfloor \frac{{}^u s_N}{\Delta_{NN}} \right\rfloor - 1 \right. \right\} \quad (5.23)$$

Since  ${}^{\ell} s_N \geq 0, {}^u s_N > 0$ , there exists the possibility when,

$$\left\lceil \frac{{}^{\ell} s_N}{\Delta_{NN}} \right\rceil = \left\lfloor \frac{{}^u s_N}{\Delta_{NN}} \right\rfloor \in \mathbb{Z}$$

In this case, there does not exist an integer,  ${}^u z_{1+N}$  that fulfills inequality (5.22) i.e.  ${}^{\beta} \mathcal{V}_1(\ell \mathbf{z}_s) = \emptyset$ .

- For  $i = N - 1$ ,

$$\zeta_{s,N-1} = \Delta_{N-1,N-1} {}^u z_N + \Delta_{N-1,N} {}^u z_{1+N} \in \mathbb{R}$$

To have  $\zeta_{s,N-1} \in \mathbb{R}$  in the range of  ${}^{\ell} s_{N-1} \leq \zeta_{s,N-1} < {}^u s_{N-1}$  means that  ${}^u z_N \in \mathbb{Z}$  is

in the range of:

$$\left\lceil \frac{\ell s_{N-1} - \Delta_{N-1,N} u z_{N+1}}{\Delta_{N-1,N-1}} \right\rceil \leq u z_N \leq \left\lfloor \frac{u s_{N-1} - \Delta_{N-1,N} u z_{N+1}}{\Delta_{N-1,N-1}} \right\rfloor - 1 \quad (5.24)$$

Let,

$$\beta \mathcal{V}_2(\ell \mathbf{z}_s) = \left\{ (u z_N, u z_{1+N}) \left| \begin{array}{l} \forall u z_{1+N} \in \beta \mathcal{V}_1(\ell \mathbf{z}_s), \\ u z_N \text{ fulfills (5.24)} \end{array} \right. \right\} \subset \mathbb{Z} \times \mathbb{Z}$$

- In general, for  $i < N$ , to fulfill  $\ell s_i \leq \zeta_{s,i} < u s_i$ , we need  $u z_{i+1} \in \mathbb{Z}$  to be within the range of:

$$\left\lceil \frac{\ell s_i - \sum_{j=i+1}^N \Delta_{ij} u z_{1+j}}{\Delta_{ii}} \right\rceil \leq u z_{i+1} \leq \left\lfloor \frac{u s_i - \sum_{j=i+1}^N \Delta_{ij} u z_{1+j}}{\Delta_{ii}} \right\rfloor - 1 \quad (5.25)$$

Since  $\ell s_i \geq 0, u s_i > 0$ , there exists the possibility when,

$$\left\lceil \frac{\ell s_i - \sum_{j=i+1}^N \Delta_{ij} u z_{1+j}}{\Delta_{ii}} \right\rceil = \left\lfloor \frac{u s_i - \sum_{j=i+1}^N \Delta_{ij} u z_{1+j}}{\Delta_{ii}} \right\rfloor$$

in which case there does not exist any integer  $u z_{i+1}$  that fulfills the inequality (5.25).

Let,

$$\beta \mathcal{V}_l(\ell \mathbf{z}_s) = \left\{ \left( \begin{array}{c} u z_{N+2-l}, \\ \dots, u z_{1+N} \end{array} \right) \left| \begin{array}{l} \forall \left( \begin{array}{c} u z_{N+1-l}, \\ \dots, u z_{1+N} \end{array} \right) \in \beta \mathcal{V}_{l-1}(\ell \mathbf{z}_s) \\ u z_{N+2-l} \text{ fulfills (5.25)} \end{array} \right. \right\} \quad (5.26)$$

For  $l = 2$  to  $l = N$ , recursively compute  $\beta \mathcal{V}_l(\ell \mathbf{z}_s)$  from (5.26). At  $l = N$ , we have  $\beta \mathcal{V}_N(\ell \mathbf{z}_s) \subset \mathbb{Z}^N$ .

Note that in this development:

1.  $\beta \mathcal{V}_N(\ell \mathbf{z}_s)$  can be an empty set. In this case there does not exist a lattice point  $\Delta^u \mathbf{z}_s$  in the overlapped region  $\mathcal{O}(\ell \mathbf{z}_s)$ . In fact note that for any  $l$  such that  $\beta \mathcal{V}_l(\ell \mathbf{z}_s) = \emptyset$ , the algorithm can be stopped since that will imply that  $\beta \mathcal{V}_N(\ell \mathbf{z}_s) = \emptyset$ .
2. Since  $\Delta_{ii} > 0$ , the bounds for  $u z_{1+N}$  given by (5.22) are finite values, hence  $\beta \mathcal{V}_1(\ell \mathbf{z}_s)$  from (5.23) is compact. Similarly the bounds for  $u z_{N+2-l}$  given by (5.25) are finite

values, hence  ${}^\beta\mathcal{V}_l(\ell\mathbf{z}_s)$  from (5.26) is compact for  $l = 2$  to  $l = N$ . Hence,  ${}^\beta\mathcal{V}_N(\ell\mathbf{z}_s)$  is compact.

The algorithms given in subsections 5.2.1.1 and 5.2.1.2 together with equation (5.10) enable us to determine the integer set  $\mathcal{V}_N \subset \mathbb{Z}^N$  for a given lattice,  $LAT(\mathbf{GH}_{LTM})$ . This process starts by determining the integer space  ${}^\alpha\mathcal{V}_N \subset \mathbb{Z}^N$  for a given lattice and known spatial domain,  $\mathcal{S}$ . For every  $\ell\mathbf{z}_s^* \in {}^\alpha\mathcal{V}_N \neq \emptyset$ , we then determine if the set  ${}^\beta\mathcal{V}_N(\ell\mathbf{z}_s^*)$  is non-empty. The set  ${}^\beta\mathcal{V}_N(\ell\mathbf{z}_s^*)$  non-empty implies that there exists  $\Delta^u\mathbf{z}_s, {}^u\mathbf{z}_s \in {}^\beta\mathcal{V}_N(\ell\mathbf{z}_s^*)$  in the overlapped region,  $\mathcal{O}(\ell\mathbf{z}_s^*)$ . Hence,  $\ell\mathbf{z}_s^*$  results in  $\mathcal{S} \cap (\mathcal{S} + \nabla^{\ell\mathbf{z}_s^*}) \cap LAT(\Delta) \neq \emptyset$ . From (5.7),  $\ell\mathbf{z}_s^*$  belongs to  $\mathcal{V}_N$ . Finding all  $\ell\mathbf{z}_s^*$  defines the set  $\mathcal{V}_N$ .

### 5.2.2 Check if a $TS(1)$ sampling lattice, $LAT(\mathbf{GH}_{LTM})$ is a non-aliasing sampling lattice

To determine if a  $TS(1)$  sampling lattice,  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing sampling lattice consider the polar form of  $LAT(\mathbf{GH}_{LTM})$  given by:

$$LAT(\mathbf{G}^*\mathbf{H}_{UTM}^*) \subset \mathbb{R}^{(1+N)}$$

where  $\mathbf{G}^* = \mathbf{G}^{-T}$  and  $\mathbf{H}_{UTM}^* = \mathbf{H}_{LTM}^{-T}$  (by Property 4.1). This polar lattice gives the locations where the spectral content of the temporal-spatial signal replicates. In this dissertation, the spectral support,  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  is assumed to be ellipsoidal (see equation (3.10)). The non-aliasing problem of  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  on  $LAT(\mathbf{G}^*\mathbf{H}_{UTM}^*)$  can be restated as a packing problem of a  $1 + N$  dimensional unit-ball,  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$  by scaling the polar generator matrix,  $\mathbf{G}^*\mathbf{H}_{UTM}^*$  such that packing the ellipsoidal  $\Theta(W, \mathbf{U})$  is equivalent to packing of the unit-ball  $\mathbf{B}$ . This scaling transformation is given by:

$$\mathbf{H}_{\mathbf{B}}^* = \mathbf{S}_\Theta \mathbf{G}^* \mathbf{H}_{LTM}^* \in \mathbb{R}^{(1+N) \times (1+N)} \quad (5.27)$$

where  $\mathbf{S}_\Theta$  is given in equation (4.35).

With this transformation  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  packs  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  if and only if  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  packs  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$ .

To determine if  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  packs the unit ball,  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$  consider the measure of Euclidean distance for any lattice point  $\xi_{\mathbf{B}} \in LAT(\mathbf{H}_{\mathbf{B}}^*)$  to the origin given by:

$$\|\xi_{\mathbf{B}}\|_2^2 = \mathbf{z}^T \mathbf{H}_{\mathbf{B}}^{*T} \mathbf{H}_{\mathbf{B}}^* \mathbf{z} \quad (5.28)$$

Then the nonzero minimum distance between the lattice point and the origin is given by:

$$\varphi_{min} = \min_{\xi_{\mathbf{B}} \in LAT(\mathbf{H}_{\mathbf{B}}^*)/\{\mathbf{0}\}} \|\xi_{\mathbf{B}}\|_2 \quad (5.29)$$

The packing condition for the  $1 + N$  dimensional unit ball,  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$  is now given by:

**Lemma 5.1**  *$LAT(\mathbf{H}_{\mathbf{B}}^*)$  packs the  $1 + N$  dimensional unit ball,  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$  if and only if  $\varphi_{min} \geq 2$ .*

**Proof:** ( $\Rightarrow$ )

- Suppose  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  packs  $\mathbf{B}$  but  $\varphi_{min} < 2$ .  $\varphi_{min} < 2$ , means that there exists a lattice point  $\xi_{\mathbf{B}} \in LAT(\mathbf{H}_{\mathbf{B}}^*)$ ,  $\xi_{\mathbf{B}} \neq \mathbf{0}$  with distance less than 2 unit from the origin.
- Then  $\mathbf{B} \cap (\mathbf{B} + \xi_{\mathbf{B}}) \neq \emptyset$ .  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  does not pack the unit ball  $\mathbf{B}$ , which contradicts our assumption.

( $\Leftarrow$ )

- Suppose  $\varphi_{min} \geq 2$  but  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  does not pack  $\mathbf{B}$ .  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  does not pack  $\mathbf{B}$  then  $\exists \xi_{\mathbf{B}} \in LAT(\mathbf{H}_{\mathbf{B}}^*)$ ,  $\xi_{\mathbf{B}} \neq \mathbf{0}$  such that  $\mathbf{B} \cap (\mathbf{B} + \xi_{\mathbf{B}}) \neq \emptyset$ .
- Then exists  $\mathbf{x} \in \mathbf{B}$  and  $\mathbf{x} \in \mathbf{B} + \xi_{\mathbf{B}}$ . For  $\mathbf{x} \in \mathbf{B}$ ,  $\|\mathbf{x}\|_2 < 1$ . Similarly, for  $\mathbf{x} \in \mathbf{B} + \xi_{\mathbf{B}}$ ,  $\|\mathbf{x} - \xi_{\mathbf{B}}\|_2 < 1$ .
- Now

$$\begin{aligned} \xi_{\mathbf{B}} &= \xi_{\mathbf{B}} - \mathbf{x} + \mathbf{x} \\ \|\xi_{\mathbf{B}}\|_2 &\leq \|\xi_{\mathbf{B}} - \mathbf{x}\|_2 + \|\mathbf{x}\|_2 \\ &< 1 + \|\mathbf{x}\|_2 < 2 \end{aligned}$$

- Hence  $\exists \xi_{\mathbf{B}} \neq \mathbf{0}$  such that  $\|\xi_{\mathbf{B}}\|_2 < 2$  which implies that  $\varphi_{min} < 2$ . ■

From this lemma, we can now state the main result of this section.

**Theorem 5.1**  *$LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)$  packs  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  if and only if  $\varphi_{min} \geq 2$ .*

**Proof:** Since  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  packs  $\Theta(W, \mathbf{U}) \subset \mathbb{R}^{(1+N)}$  if and only if  $LAT(\mathbf{H}_{\mathbf{B}}^*)$  packs  $\mathbf{B} \subset \mathbb{R}^{(1+N)}$ , by Lemma 5.1 we arrived at the desired result.  $\blacksquare$

### 5.2.3 Algorithm to determine if a given lattice $LAT(\mathbf{GH}_{LTM})$ is a non-aliasing $TS(1)$ sampling lattice

For a given lattice,  $LAT(\mathbf{GH}_{LTM})$ , the algorithm to find a non-aliasing  $TS(1)$  sampling lattice can be summarized as follows:

#### Algorithm 5.1

**Step 1:** For the given lattice start by constructing the integer set  $\mathcal{V}_N \subset \mathbb{Z}^N$ .

**Step 2:** If  $\mathcal{V}_N = \{\mathbf{0}\}$  then go to **Step 3** else go to **Step 4**.

**Step 3:** Compute the re-scaled polar generator matrix  $\mathbf{H}_{\mathbf{B}}^* \in \mathbb{R}^{(1+N) \times (1+N)}$  given by (5.27), then use the closest point search algorithm (see Appendix C.3) to determine  $\varphi_{min} \in \mathbb{R}$ . If  $\varphi_{min} \geq 2$ , then by Theorem 5.1  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing  $TS(1)$  sampling lattice (i.e. its polar lattice,  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  packs  $\Theta(W, \mathbf{U})$ ) and terminate else go to **Step 4**.

**Step 4:**  $LAT(\mathbf{GH}_{LTM})$  is not a non-aliasing  $TS(1)$  sampling lattice and terminate

Algorithm 5.1 enable us to check if there exists a non-aliasing  $TS(1)$  sampling lattice in the set  $\mathcal{LEP}(\mathbf{G})$  for a given grid,  $LAT(\mathbf{G}) \in \mathcal{L}_{grid}$ . Since  $\mathcal{LEP}(\mathbf{G}) \supset \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ , hence existence of a non-aliasing  $TS(1)$  sampling lattice implies that  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  is non-empty and  $LAT(\mathbf{G})$  is an admissible grid.

## 5.3 Optimization Algorithm

To solve Problem 5.1, all non-aliasing  $TS(1)$  sampling lattice for all possible grids need to be characterized (see equation (5.5)). The previous section gives the characterization of the non-aliasing  $TS(1)$  sampling lattices for a given grid. Each grid,  $LAT(\mathbf{G})$  is parameterized by the triple  $(T, \mathbf{A}, \mathbf{M})$ . These parameters are given as follows:

1.  $\mathbf{A} \in \mathbb{R}^N$  is the maximum  $N$ -dimensional spatial range and is known in advance for the optimal  $TS(1)$  sampling design.

2. The number of discretized spatial points at each  $i$ -th spatial dimension  $M_i \in \mathbb{Z}^+$  belongs to the finite  $N$  dimensional integer set,

$${}^1\mathcal{V}_{\mathbf{M}} = [\mathbf{M}_{min}, \mathbf{M}_{max}] \subset \mathbb{Z}^N \quad (5.30)$$

where

- (a)  $\mathbf{M}_{min} = \lfloor 2\mathbf{A} \cdot * \mathbf{U} \rfloor \in \mathbb{Z}^N$  by the spatial Nyquist condition given in (3.30), and
- (b)  $\mathbf{M}_{max} = \lceil \mathbf{A} \cdot / \varepsilon \rceil \in \mathbb{Z}^N$  with  $\varepsilon$  denoting the sensor resolution.

3. The inter-sampling time  $T \in \mathbb{R}^+$  at a given  $\mathbf{M} \in \mathbb{Z}^N$  lies in the bounds of

$$T_{min}(\mathbf{M}) \leq T \leq T_{max}(\mathbf{M}) \quad (5.31)$$

where

- (a) By the temporal-spatial Nyquist conditions given by (3.31) and (3.30), the minimum inter-sampling time for a given  $\mathbf{M}$ ,  $T_{min}(\mathbf{M})$  is:

$$T_{min}(\mathbf{M}) = 1 / \left( 4W \prod_{i=1}^N A_i U_i \right) \quad (5.32)$$

- (b)  $T_{max}(\mathbf{M})$  is obtained from the statement of Corollary 4.1 and inequality (4.39). From Corollary 4.1, for a lattice to be a  $TS(1)$  sampling lattice,  $\prod_{i=2}^{1+N} \ell h_{ii} \geq \min(\mathbf{M})$ . Hence, from inequality (4.39):

$$T \leq \frac{\text{prod}(\mathbf{M} \cdot / \mathbf{A})}{\min(\mathbf{M}) \vartheta[LAT(\mathbf{V}_{crit}^{\Theta})]} := T_{max}(\mathbf{M}) \quad (5.33)$$

The real vector  $\mathbf{A} \in \mathbb{R}^N$  is a known constant, the set of possible integer vector  $\mathbf{M}$  that discretized the spatial domain,  $\mathcal{V}_{\mathbf{M}}$  is compact and the inter-sampling time,  $T$  is upper and lower bounded for a given  $\mathbf{M}$ . Hence, an optimization algorithm can now be proposed to solve Problem 5.1, given as follows:

### Algorithm 5.2

**Step 1:** Start by taking  $\mathbf{M} \leftarrow \mathbf{M}_{min}$ ,  $T \leftarrow T_{max}(\mathbf{M})$ . Hence, we can define the grid,  $LAT(\mathbf{G})$  and the set  $\mathcal{LEP}(\mathbf{G})$  by (4.42).

---

<sup>1</sup>For example, for  $N = 2$ , given that  $\mathbf{M}_{min} = (1, 1)$  and  $\mathbf{M}_{max} = (3, 2)$ , then

$$\mathcal{V}_{\mathbf{M}} = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (3, 2)\} \subset \mathbb{Z} \times \mathbb{Z}$$

**Step 2:** Find if  $\exists \text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  such that  $\text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  by Algorithm 5.1. If yes, then go to **Step 3**, otherwise reduce  $T$  by  $\Delta T$  where  $\Delta T < T_{max}(\mathbf{M}) - T_{min}(\mathbf{M})$  pre-selected arbitrarily, update the grid,  $\text{LAT}(\mathbf{G})$  and set  $\mathcal{LEP}(\mathbf{G})$  by (4.42). Then repeat **Step 2**.

**Step 3:** Initialize bisection on  $T$  by taking  $T_u \leftarrow T + \Delta T$ ,  $T_l \leftarrow T$ .

**Step 4:** Do bisection. Take  $T \leftarrow (T_u + T_l)/2$ , update the grid,  $\text{LAT}(\mathbf{G})$  and set  $\mathcal{LEP}(\mathbf{G})$  by (4.42) then find if  $\exists \text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  such that  $\text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  by Algorithm 5.1. If yes, take  $T_l \leftarrow T$ , otherwise take  $T_u = T$ . Repeat **Step 4** until  $T_u - T_l < \delta_{tol}$  where  $\delta_{tol}$  gives the desired tolerance, then set  $T^* \leftarrow T$ ,  $\mathbf{M}^* \leftarrow \mathbf{M}$ ,  $\mathbf{G}^* \mathbf{H}_{LTM}^* \leftarrow \mathbf{GH}_{LTM}$  and go to **Step 5**.

**Step 5:** Go to the next  $N$ -length integer vector enumeration,  $\mathbf{M}$  of the set  $\mathcal{V}_{\mathbf{M}}$  given by (5.30). If all the elements in  $\mathcal{V}_{\mathbf{M}}$  are exhausted go to **Step 6**, else update  $\text{LAT}(\mathbf{G})$  and  $\mathcal{LEP}(\mathbf{G})$  and find if  $\exists \text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  such that  $\text{LAT}(\mathbf{GH}_{LTM}) \in \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  by Algorithm 5.1. If not, repeat **Step 5**. If yes then take  $T_u \leftarrow T_{max}(\mathbf{M})$ ,  $T_l \leftarrow T$  and go to **Step 4**.

**Step 6:** Terminate optimization. We have the optimal  $\mathbf{M}$  and  $T$  given by  $\mathbf{M}^*$  and  $T^*$  respectively and the optimal  $TS(1)$  sampling sequence,  $\boldsymbol{\alpha}^*(k)$  is obtained from the lattice,  $\text{LAT}(\mathbf{H}_{LTM}^*)$ .

Bisection on  $T \in \mathbb{R}^+$  at fixed  $\mathbf{M} \in \mathbb{Z}^N$  (**Step 4**) is possible due to the following proposition:

**Proposition 5.1** *If*

(i) *there does not exist a lattice  $\text{LAT}(\mathbf{G}_1 \mathbf{H}_{LTM})$  in the set  $\mathcal{LEP}(\mathbf{G}_1)$  where  $\mathbf{G}_1 = \text{diag}(T_1, \mathbf{A} \cdot / \mathbf{M})$  such that  $\text{LAT}(\mathbf{G}_1 \mathbf{H}_{LTM})$  gives a non-aliasing  $TS(1)$  sampling lattice and,*

(ii) *given the set  $\mathcal{LEP}(\mathbf{G}_2)$  where  $\mathbf{G}_2 = \text{diag}(T_2, \mathbf{A} \cdot / \mathbf{M})$  with  $T_2 > T_1$*

*then there does not exist any lattice,  $\text{LAT}(\mathbf{G}_2 \mathbf{H}_{LTM})$  in the set  $\mathcal{LEP}(\mathbf{G}_2)$  such that  $\text{LAT}(\mathbf{G}_2 \mathbf{H}_{LTM})$  gives a non-aliasing  $TS(1)$  sampling lattice.*

**Proof:**

- Let define the following sets of lower-triangular integer HNF matrices:

$$\mathcal{H}(\mathbf{G}_1) = \{\mathbf{H}_{LTM} | LAT(\mathbf{G}_1 \mathbf{H}_{LTM}) \in \mathcal{LEP}(\mathbf{G}_1)\} \quad (5.34)$$

$$\mathcal{H}(\mathbf{G}_2) = \{\mathbf{H}_{LTM} | LAT(\mathbf{G}_2 \mathbf{H}_{LTM}) \in \mathcal{LEP}(\mathbf{G}_2)\} \quad (5.35)$$

- Since the lower-triangular integer HNF matrix encodes the sampling sequence, the set  $\mathcal{H}(\mathbf{G}_1)$  and  $\mathcal{H}(\mathbf{G}_2)$  give all possible sampling sequences (patterns) to be evaluated in the optimization algorithm.
- From (4.42), the diagonal components of the lower-triangular HNF matrix,  $\mathbf{H}_{LTM}$  i.e.  $(\ell h_{22}, \dots, \ell h_{1+N,1+N})$  in the set  $\mathcal{LEP}(\mathbf{G}_1)$  fulfills the following inequality:

$$\prod_{i=2}^{1+N} \ell h_{ii} \leq \frac{1}{T_1 \vartheta[LAT(\mathbf{V}_{crit}^\ominus)]} \prod_{i=1}^N \frac{M_i}{A_i} \quad (5.36)$$

Similarly the diagonal components of the lower-triangular HNF matrix,  $\mathbf{H}_{LTM}$  i.e.  $(\ell h_{22}, \dots, \ell h_{1+N,1+N})$  in the set  $\mathcal{LEP}(\mathbf{G}_2)$  fulfills the following inequality:

$$\prod_{i=2}^{1+N} \ell h_{ii} \leq \frac{1}{T_2 \vartheta[LAT(\mathbf{V}_{crit}^\ominus)]} \prod_{i=1}^N \frac{M_i}{A_i} \quad (5.37)$$

- Since  $T_2 > T_1$ , more  $\mathbf{H}_{LTM}$  diagonal components combinations can be achieved to fulfill inequality (5.36) than for inequality (5.37).
- By the HNF theorem (Theorem 4.2),

$$\mathcal{H}(\mathbf{G}_2) \subset \mathcal{H}(\mathbf{G}_1) \quad (5.38)$$

This implies that sampling sequences found in  $\mathcal{LEP}(\mathbf{G}_2)$  can also be found in  $\mathcal{LEP}(\mathbf{G}_1)$ .

- Also for  $T_2 > T_1$  and for a given  $\mathbf{H}_{LTM}$  matrix, the polar lattice  $LAT(\mathbf{G}_2^{-T} \mathbf{H}_{LTM}^{-T})$  (see lattice Property 4.1) have lattice points that are closer to one another than the polar lattice  $LAT(\mathbf{G}_1^{-T} \mathbf{H}_{LTM}^{-T})$ .
- Hence sampling sequences found in  $\mathcal{LEP}(\mathbf{G}_2)$  can also similarly be found in  $\mathcal{LEP}(\mathbf{G}_1)$  and sampling lattices in  $\mathcal{LEP}(\mathbf{G}_2)$  yields frequency content packing that is tighter than sampling lattices in  $\mathcal{LEP}(\mathbf{G}_1)$ .
- Hence if there exists no non-aliasing  $TS(1)$  sampling lattice in  $\mathcal{LEP}(\mathbf{G}_1)$ , we will not find any non-aliasing  $TS(1)$  sampling lattice in  $\mathcal{LEP}(\mathbf{G}_2)$ .



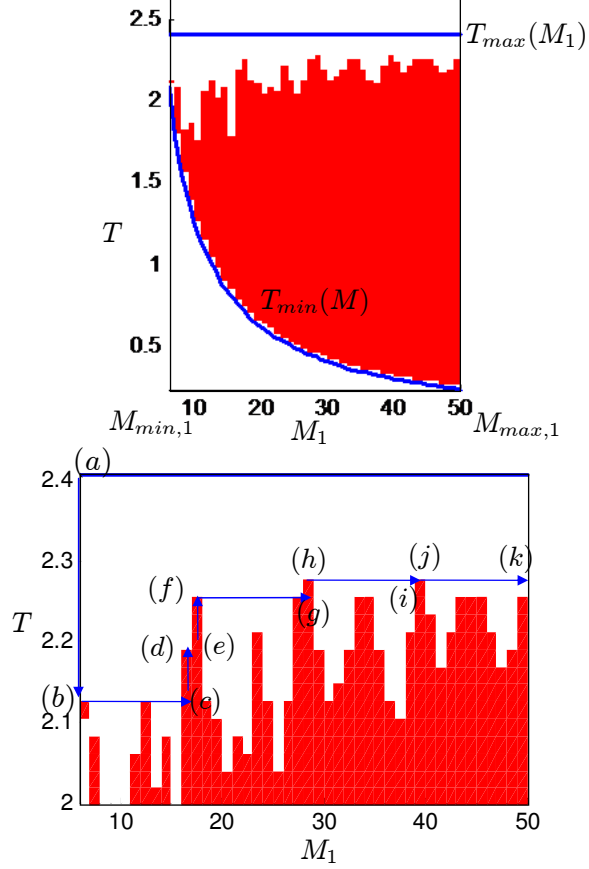


Figure 5.3: An example of the  $(T, M_1)$  optimization search space for the case of the TRC ( $1 + N = 2$ ). The shaded area shows the area where there exists  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  such that  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  i.e.  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  is non-empty. ■

### 5.3.1 Illustrated example of optimization process

Figure 5.3 shows an example of the  $(T, M_1)$  optimization search space for the case of the time-varying TRC ( $1 + N = 2$ ). The shaded area shows the area where there exists  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  such that  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  i.e.  $\mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$  is non-empty. Based on this figure, the optimization algorithm will proceed as follows:

1. Start at (a) [**Step 1**]
2. (a)  $\rightarrow$  (b) [**Step 2**  $\rightarrow$  **Step 3**  $\rightarrow$  **Step 4**]

3.  $(b) \rightarrow (c)$  [repeat **Step 5**]
4.  $(c) \rightarrow (d) \rightarrow (e) \rightarrow (f)$  [**Step 4**  $\rightarrow$  **Step 5**  $\rightarrow$  **Step 4**]
5.  $(f) \rightarrow (g)$  [repeat **Step 5**]
6.  $(g) \rightarrow (h)$  [**Step 4**  $\rightarrow$  **Step 5**  $\rightarrow$  **Step 4**]
7.  $(h) \rightarrow (i)$  [repeat **Step 5**]
8.  $(i) \rightarrow (j)$  [**Step 4**  $\rightarrow$  **Step 5**  $\rightarrow$  **Step 4**]
9.  $(j) \rightarrow (k)$  [repeat **Step 5**]
10. Stop at  $(k)$  [**Step 6**]
11. Optimized parameters found at  $(j)$

At a given grid,  $LAT(\mathbf{G})$ , the cardinality of the set of all feasible non-aliasing  $TS(1)$  sampling lattices,  $\mathcal{LEP}(\mathbf{G})$  can be very large. Hence, to reduce the computational complexity associated with this optimization algorithm, we introduce lattice restriction conditions as given in the next section.

## 5.4 Lattice Restriction Conditions

In this section, the lattice restriction conditions are given to reduce the computational complexity at **Step 2**, **Step 4** and **Step 5** of the optimization Algorithm 5.2.

### 5.4.1 Restriction due to the loose packing of the frequency content replicas

From (4.42), possibly non-aliasing  $TS(1)$  sampling lattice,  $LAT(\mathbf{GH}_{LTM}) \in \mathcal{LEP}(\mathbf{G})$  on the grid,  $LAT(\mathbf{G})$  requires that the diagonal components of the integer matrix  $\mathbf{H}_{LTM}$  matrix take the values of:

- (i)  ${}^\ell h_{ij} = 0$ , for  $i < j$
- (ii)  ${}^\ell h_{11} = 1$
- (iii)  ${}^\ell h_{ii} > 0, \forall i \in \{2, \dots, 1 + N\}$
- (iv)  ${}^\ell h_{ij} \leq 0$  and  $|{}^\ell h_{ij}| < {}^\ell h_{ii}, i > j$
- (v)  ${}^\ell h_{1+N,1+N} \geq M_N$

(vi)  $({}^\ell h_{22}, \dots, {}^\ell h_{1+N, 1+N})$  fulfills the inequality (given by (4.39))

$$\prod_{i=2}^{1+N} {}^\ell h_{ii} \leq \frac{1}{T\vartheta[LAT(\mathbf{V}_{crit}^\Theta)]} \prod_{i=1}^N \frac{M_i}{A_i}$$

However we will show here that lattices with

$${}^\ell h_{ii} > M_{i-1}, i = 2, \dots, 1 + N \quad (5.39)$$

resulted in loose packing in the frequency content replicas. Hence lattices with this property can be ignored in the optimization process because they will not be selected as the optimal  $TS(1)$  sampling lattice.

To formulate this condition, let

$$\mathcal{LE}_r(\mathbf{G}) = \left\{ LAT(\mathbf{G}\mathbf{H}_{LTM}) \left| \begin{array}{l} \mathbf{H}_{LTM} \text{ given by:} \\ \text{(i) } {}^\ell h_{ij} = 0, \text{ for } i < j, \\ \text{(ii) } {}^\ell h_{11} = 1 \\ \text{(iii) } {}^\ell h_{ii} > M_{i-1}, \forall i = 2, \dots, 1 + N \\ \text{(iv) } {}^\ell h_{ij} \leq 0 \text{ and } |{}^\ell h_{ij}| < {}^\ell h_{ii}, i > j \end{array} \right. \right\} \quad (5.40)$$

be the set of all restricted sampling lattices with property given by (5.39).

By the definition of the set  $\mathcal{LE}_s(\mathbf{G})$  from (4.24),  $\mathcal{LE}_r(\mathbf{G}) \subset \mathcal{LE}_s(\mathbf{G})$ . From Corollary 4.1,  $\mathcal{LE}_s(\mathbf{G}) \subseteq \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S})$ . Hence,  $\mathcal{LE}_r(\mathbf{G}) \cap \mathcal{P}(\Theta) \subset \mathcal{L}(\mathbf{G}) \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ . Let,

$$\mathcal{LEP}_r(\Theta) = \bigcup_{LAT(\mathbf{G}) \in \mathcal{L}_{grid}} \mathcal{LE}_r(\mathbf{G}) \cap \mathcal{P}(\Theta) \quad (5.41)$$

Note that,  $\mathcal{LEP}_r(\Theta) \subset \mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ . The lattice restriction condition can now be stated as follows:

**Proposition 5.2** *If  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  is an optimal  $TS(1)$  sampling lattice then the lattice,  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  belongs to the set  $(\mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)) / \mathcal{LEP}_r(\Theta)$ .*

**Proof:**

- Suppose  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  is an optimal  $TS(1)$  sampling lattice but  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*) \in \mathcal{LEP}_r(\Theta)$ . Since,  $\mathcal{LEP}_r(\Theta) \subset \mathcal{L} \cap \mathcal{E}(1, \mathcal{S}) \cap \mathcal{P}(\Theta)$ ,  $LAT(\mathbf{G}^*\mathbf{H}_{LTM}^*)$  is a  $TS(1)$  sampling lattice.

- From (4.16),

$$\mathbf{G}^* \mathbf{H}_{LTM}^* = \begin{pmatrix} T^* & 0 \\ \mathbf{s}^* & \nabla^* \end{pmatrix} \in \mathbb{R}^{(1+N) \times (1+N)}$$

- Hence, by Theorem 4.3,

$$\mathcal{S} \cap (\mathcal{S} + \mathbf{b}) \cap LAT(\Delta^*) = \emptyset, \quad \forall \mathbf{b} \in LAT(\nabla^*) \subset \mathbb{R}^N, \mathbf{b} \neq \mathbf{0} \quad (5.42)$$

- From (5.40) and (5.41)

$$\ell h_{ii}^* > M_{i-1}^*, \quad \forall i \in \{2, \dots, 1+N\}$$

- Then the non-overlapping spatial domain,  $\mathcal{S} \subset \mathbb{R}^{(1+N)}$  replicated at  $\mathbf{b} \in LAT(\nabla^*)$ ,  $\mathbf{b} \neq \mathbf{0}$  are loosely pack. There always exists gaps in between the non-overlapping replicas of  $\mathcal{S} \subset \mathbb{R}^{(1+N)}$  on  $LAT(\nabla^*)$ . Hence it is still possible to further constrict  $LAT(\nabla^*)$  while satisfying (5.42).
- By Property 4.1 constricting  $LAT(\nabla^*)$  results in expansion in the spatial frequency domain, widening the gaps of replicas of the compact frequency content,  $\Theta(W, \mathbf{U})$  in sampled signal's frequency domain. Further tradeoff of spatial frequency for temporal frequency is possible. Therefore the inter-sampling time can still be possibly increased.
- Since we cannot do worst than the original non-constricted  $TS(1)$  sampling lattice, a better  $TS(1)$  lattice can be found as the gaps in between the non-overlapping replicas of  $\mathcal{S}$  on  $LAT(\nabla^*)$  gets smaller. Therefore  $LAT(\mathbf{G}^* \mathbf{H}_{LTM}^*)$  is not an optimal  $TS(1)$  lattice, contradicting our assumption. ■

#### 5.4.2 Restriction due to the equivalence of a lattice to its reflected lattice in the temporal-spatial frequency domain

Let  $LAT(\mathbf{V}_1^*)$  be the polar lattice of the sampling lattice  $LAT(\mathbf{G}\mathbf{H}_{LTM_1})$ . Let  $LAT(\mathbf{V}_2^*)$  be the reflected polar lattice of  $LAT(\mathbf{V}_1^*)$  with  $\mathbf{V}_2^*$  defined by:

$$\mathbf{V}_2^* = \mathbf{R}\mathbf{V}_1^* \mathbf{Q} \quad (5.43)$$

where  $\mathbf{R} \in \mathbb{Z}^{(1+N) \times (1+N)}$  is a reflection matrix i.e.  $\mathbf{R}^T \mathbf{R} = \mathbf{R}\mathbf{R}^T = \mathbf{I}$ ,  $\mathbf{R} \neq \mathbf{I}$ ,  $\mathbf{R}$  is a diagonal matrix and  $\mathbf{Q}$  is a unimodular matrix. Finally, let  $LAT(\mathbf{G}\mathbf{H}_{LTM_2})$  denotes the

sampling lattice for the polar lattice  $LAT(\mathbf{V}_1^*)$ .

Note that  $LAT(\mathbf{V}_1^*) = LAT(\mathbf{V}_2^*)$ . This is arrived by observing that  $\mathbf{R}$  is unimodular and diagonal. Hence, we have  $\mathbf{V}_2^* = \mathbf{R}\mathbf{V}_1^*\mathbf{Q} = \mathbf{V}_1^*\mathbf{R}\mathbf{Q}$  with  $\mathbf{R}\mathbf{Q}$  unimodular. By lattice Property 4.2  $LAT(\mathbf{V}_1^*) = LAT(\mathbf{V}_2^*)$ . Hence if the sampling lattice  $LAT(\mathbf{G}\mathbf{H}_{LTM_1})$  has been evaluated, the sampling lattice  $LAT(\mathbf{G}\mathbf{H}_{LTM_2})$  can be ignored by the fact that both yield the same frequency content packing (since  $LAT(\mathbf{V}_1^*) = LAT(\mathbf{V}_2^*)$ ).

The idea then is that for a given sampling lattice being evaluated,  $LAT(\mathbf{G}\mathbf{H}_{LTM_1})$  we find the reflected polar lattice with generator matrix defined by (5.43) and ignored the sampling lattice,  $LAT(\mathbf{G}\mathbf{H}_{LTM_2})$  in the optimization process. To find  $LAT(\mathbf{G}\mathbf{H}_{LTM_2})$ , by lattice Property 4.1, equation (5.43) can be rewritten as:

$$\begin{aligned} (\mathbf{G}\mathbf{H}_{LTM_2})^{-T} &= \mathbf{R}(\mathbf{G}\mathbf{H}_{LTM_1})^{-T}\mathbf{Q} \\ \mathbf{H}_{LTM_2} &= \mathbf{R}\mathbf{H}_{LTM_1}\mathbf{Q}^{-T} \end{aligned} \quad (5.44)$$

where  $\mathbf{Q}^{-T}$  is still unimodular. Therefore, the reflected lattice  $LAT(\mathbf{G}\mathbf{H}_{LTM_2})$  can be found by applying the column-reduced HNF algorithm in Appendix C.2, on the matrix  $\mathbf{R}\mathbf{H}_{LTM_1}$  to find the  $\mathbf{H}_{LTM_2}$  matrix. However, the HNF algorithm incurs a computational penalty if used directly in Algorithm 5.2. To avoid this penalty, an ordered logical list can be deduced off-line that indicate reflected matrices that can be ignored in the optimization process for a given grid.

## 5.5 Design of $TS(n)$ Sampling Lattices

The idea in designing the  $TS(n)$ ,  $n > 1$  sampling lattice with respect to the spatial domain,  $\mathcal{S} \subset \mathbb{R}^N$  is to construct it by extending the optimal  $TS(1)$  sampling lattice with respect to a smaller spatial domain,  $\mathcal{S}_\top \subset \mathcal{S}$ . This is illustrated in Figure 5.4 for  $TS(2)$  and  $N = 1$  by simply tiling  $\mathcal{S}_\top$  uniformly in  $\mathcal{S}$ .

In general to uniformly tile  $\mathcal{S}_\top \subset \mathcal{S}$  in  $\mathcal{S}$  we have:

$$\mathcal{S}_\top = \prod_{i=1}^N \left[ 0, \frac{A_i}{\rho_i} \right) \subset \mathcal{S} \quad (5.45)$$

where the  $N$ -tuples integer vector,  $\boldsymbol{\rho} = [\rho_1, \dots, \rho_N] > \mathbf{0}$  is selected such that  $n, \mathcal{S}_\top \subset \mathcal{S}$  uniformly tiled  $\mathcal{S}$  i.e.

$$\mathcal{S} = \bigcup_{j=1}^n (\mathcal{S}_\top + \mathbf{g}_j) \quad (5.46)$$

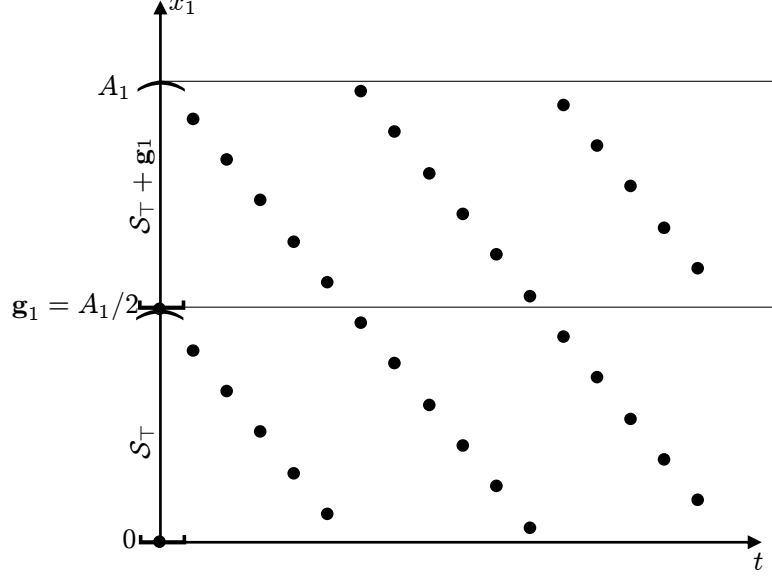


Figure 5.4: An example of  $TS(2)$  sampling with respect to  $\mathcal{S} = [0, A_1)$  by extending the  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top$ . In this case note that we have 2  $\mathcal{S}_\top$  uniformly tiled  $\mathcal{S}$  i.e.  $\mathcal{S} = \mathcal{S}_\top \cup (\mathcal{S} + \mathbf{g}_1)$ .

where  $\mathbf{g} = [g_{1j}, \dots, g_{Nj}]^T$  and  $g_{ij} = A_i p_{ij} / \rho_i, i = 1, \dots, N$  with  $\mathbf{p}_j = [p_{1j}, \dots, p_{Nj}]^T \in [\mathbf{0}, \boldsymbol{\rho} - \mathbf{1}] \subset \mathbb{Z}^N, j = 1, \dots, n$ . An example of this uniform tiling is shown in Figure 5.5 for the case of  $N = 2$  spatial dimension. Based on this construction, the  $TS(n), n > 1$  sampling lattice can be constructed from  $TS(1)$  sampling lattice as stated by the following proposition:

**Proposition 5.3** *If  $LAT(\mathbf{GH}_{LTM})$  is a  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top$  then  $LAT(\mathbf{GH}_{LTM})$  with respect to  $\mathcal{S}$  is a  $TS(n), n > 1$  sampling lattice.*

**Proof:**

- By Lemma 4.1 if  $LAT(\mathbf{GH}_{LTM}), \mathbf{GH}_{LTM} = \begin{pmatrix} T & 0 \\ \mathbf{s} & \nabla \end{pmatrix}$  is a  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top$ , then

$$|\mathcal{S}_\top \cap (k\mathbf{s} + LAT(\nabla))| \leq 1, \forall k \in \mathbb{Z}^+ \quad (5.47)$$

- Since  $\mathcal{S}_\top$  uniformly tiled  $\mathcal{S}$  as given by (5.46), hence for all  $j$

$$(LAT(\nabla) + \mathbf{g}_j) = LAT(\nabla) \quad (5.48)$$

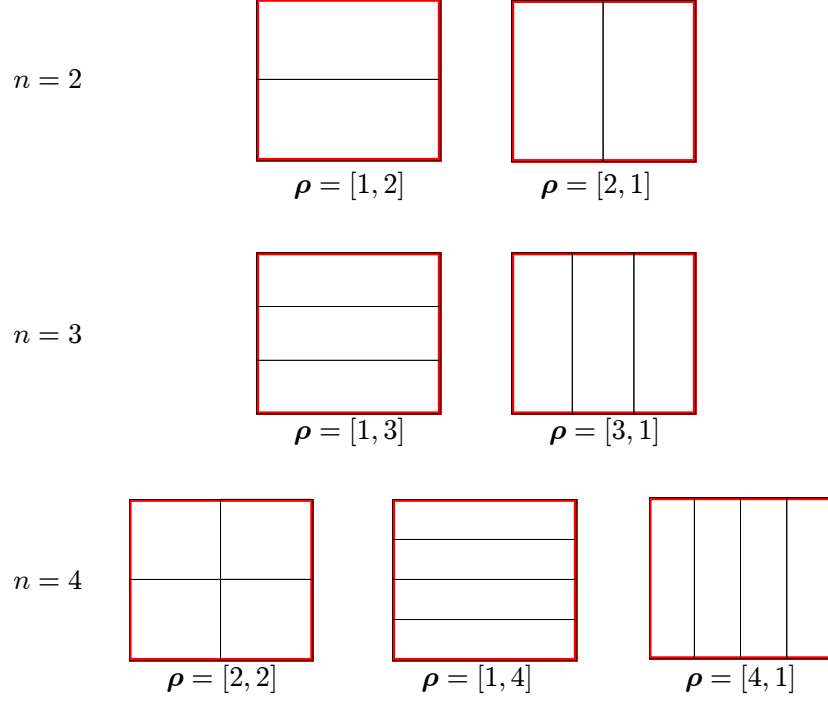


Figure 5.5: Uniform tiling of the spatial domain,  $\mathcal{S}$  for  $N = 2$

and

$$\left| (\mathcal{S}_\top + \mathbf{g}_j) \cap (k\mathbf{s} + LAT(\nabla) + \mathbf{g}_j) \right| \leq 1, \forall k \in \mathbb{Z}^+ \quad (5.49)$$

- Hence, from (5.48) and (5.47), for all  $j$

$$\left| (\mathcal{S}_\top + \mathbf{g}_j) \cap (k\mathbf{s} + LAT(\nabla)) \right| \leq 1, \forall k \in \mathbb{Z}^+ \quad (5.50)$$

- For,

$$\begin{aligned} \left| \mathcal{S} \cap (k\mathbf{s} + LAT(\nabla)) \right| &= \left| \bigcup_{j=1}^n (\mathcal{S}_\top + \mathbf{g}_j) \cap (k\mathbf{s} + LAT(\nabla)) \right| \\ &= \left| (\mathcal{S}_\top + \mathbf{g}_1) \cap (k\mathbf{s} + LAT(\nabla)) \right| + \\ &\quad \left| (\mathcal{S}_\top + \mathbf{g}_2) \cap (k\mathbf{s} + LAT(\nabla)) \right| + \dots + \\ &\quad \left| (\mathcal{S}_\top + \mathbf{g}_n) \cap (k\mathbf{s} + LAT(\nabla)) \right| \end{aligned}$$

- From (5.50),

$$\left| \mathcal{S} \cap (k\mathbf{s} + LAT(\nabla)) \right| \leq n, \forall k \in \mathbb{Z}^+ \quad (5.51)$$

- Hence, by Lemma 4.1,  $LAT(\mathbf{GH}_{LTM})$  is a  $TS(n)$  sampling lattice with respect to  $\mathcal{S}$ . ■

Hence by Proposition 5.3, an optimal  $TS(1)$  sampling lattice designed with respect to  $\mathcal{S}_\top$  will give a  $TS(n), n > 1$  sampling lattice with respect to  $\mathcal{S}$ . Now, under a signal uniformity condition it can be shown that the resulting  $TS(n)$  sampling is a non-aliasing sampling lattice. Let,

$$\omega_\top(t, \mathbf{x}) = \begin{cases} \omega(t, \mathbf{x}) & \text{for } \mathbf{x} \in \mathcal{S}_\top \\ 0 & \text{otherwise} \end{cases} \quad (5.52)$$

be the signal segment within the small spatial domain,  $\mathcal{S}_\top \subset \mathcal{S}$ . Let  $\Omega_\top(f, \mathbf{u})$  denotes the power spectral density of  $\omega_\top(t, \mathbf{x})$ . By the signal uniformity condition we assumed that:

$$\Omega(f, \mathbf{u}) = \Omega_\top(f, \mathbf{u}), \quad \forall (f, \mathbf{u}) \in \mathbb{R} \times \mathbb{R}^N \quad (5.53)$$

Since  $\Omega(f, \mathbf{u})$  is essentially bandlimited to a closed support region,  $\Theta$  as given by (3.8) ( $\Theta$  assumed to be an ellipsoid in this dissertation), by (5.53),  $\Omega(f, \mathbf{u})$  is also essentially bandlimited to a closed support region,  $\Theta$  i.e.

$$\int_{\mathbb{R} \times \mathbb{R}^N / \Theta} \Omega_\top(f, \mathbf{u}) df d\mathbf{u} \leq \epsilon$$

This enables us to make the following conclusion:

**Proposition 5.4** *If  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top \subset \mathcal{S}$  ( $\mathcal{S}$  defined by (5.45)) and the signal is uniform then  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing  $TS(n), n > 1$  sampling lattice with respect to  $\mathcal{S}$*

**Proof:**

- if  $LAT(\mathbf{GH}_{LTM})$  is a non-aliasing  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top$  then  $\Theta$  packs its polar lattice,  $LAT(\mathbf{V}^*)$ ,  $\mathbf{V}^* = (\mathbf{GH}_{LTM})^{-T}$  (by lattice Property 4.1) i.e.

$$(\Theta + \boldsymbol{\xi}_1^*) \cap (\Theta + \boldsymbol{\xi}_2^*) \neq \emptyset, \forall \boldsymbol{\xi}_1^*, \boldsymbol{\xi}_2^* \in LAT(\mathbf{V}^*), \boldsymbol{\xi}_1^* \neq \boldsymbol{\xi}_2^* \quad (5.54)$$

- By Proposition 5.3  $LAT(\mathbf{GH}_{LTM})$  with respect to  $\mathcal{S}$  gives the  $TS(n), n > 1$  sampling lattice.
- Also under the signal uniformity condition, condition given by (5.53) holds.



- Hence, the packing condition given by (5.54) also holds for  $TS(n), n > 1$  sampling lattice.

■

The sufficiency of Proposition 5.3 in defining  $TS(n), n > 1$  sampling lattice with respect to  $\mathcal{S}$  using  $TS(1)$  sampling lattice with respect to  $\mathcal{S}_\top$  means that there exists  $TS(n), n > 1$  sampling lattice that is not defined by such construction. Hence, while the resulting  $TS(n), n > 1$  sampling design constructed with the optimal  $TS(1)$  sampling design with respect to  $\mathcal{S}_\top$  does not result in aliasing (by Proposition 5.4), it is *not* optimal i.e. it does not necessarily give the largest inter-sampling time,  $T$ . Nonetheless the  $TS(n), n > 1$  sampling design allows us to arrive at a compromise between the  $n$  sensors hardware and inter-sampling time,  $T$  without loss in performance. This is of significant practical importance because typically the inter-sampling time of physical system is fixed e.g. the print cycle time for printing system is fixed.

## 5.6 Application of $TS(n)$ Sampling Design

In this section, the  $TS(n)$  sampling design is applied for sampling the tone reproduction curve (TRC) and the color reproduction characteristics (CRC) function. Currently, process sensing for xerography consists of printing small tone patches in the unused areas of the photoreceptor and measuring their densities using the enhanced toner-area coverage (ETAC) sensors<sup>2</sup>. Typically, only 3 to 5 patches are printed every few photoreceptor belt cycles. This amounts to sampling either the TRC 3 to 5 points once in a while. Increasing the number of test patches increase hardware needs as well as consumable (toner) and productivity. While this is the current state of art, improvement in sensor technologies has allowed for the use of a spectrophotometer<sup>3</sup> for inline sensing directly on paper. An example of such a system is the X-Rite vipPAQ [55]. Full color quality measurement (i.e.  $L^*a^*b^*$  values instead of just intensity with the ETAC sensor) can be measured. This new physical sensing approach is used in this work. While the three dimensional  $L^*a^*b^*$  is measured, for the case of the TRC a one dimensional output tone value suffices. A parameterizations approach proposed in [28] is used to convert the measured  $L^*a^*b^*$  to the one dimensional representation. This technique is found to be less noisy, an added benefit of using the spectrophotometer's measurement. However even with this new sensing technique we are still limited to printing and measuring small number of  $n$  samples at each print-cycle, which motivates the need

<sup>2</sup>The ETAC sensor utilizes illumination from a single LED which has been tuned to a specific wavelength and a PIN diode as the detector[8].

<sup>3</sup>A spectrophotometer is a device that can measure spectral reflectance or transmittance of color as a function of the wavelength. This measurement can then be converted to  $L^*a^*b^*$  values for a standard lighting and observer [18].

for the  $TS(n)$  sampling design. Using this sampling approach we are able to obtain the reconstructed time-varying TRC and CRC using a periodic Kalman filter as presented in section 3.3.

### 5.6.1 $TS(n)$ sampling design for TRC

In this design the ellipsoidal temporal-spatial compact spectral support of the time-varying TRC,  $\Theta(W, U_1)$  is parameterized by  $W = 0.04$  [Hz],  $U_1 = 3.0$  [cycles/tone] as shown in Figure 3.1. These parameters are taken based on the frequency content analysis on a collection of time-varying TRC,  $\bar{\omega}_1(k) \in \mathbb{R}^{M_1=9}$  collected from  $k = 0$  to  $k = 357$ .  $\Theta(W, U_1)$  is found by wrapping an ellipse on the most significant power spectral density with procedure in Appendix B.2. In addition the maximum spatial range,  $A_1 = 1$  where  $\mathbf{tone} = 1$  corresponds to a solid tone. The maximum sensor resolution,  $\varepsilon_1 = 0.02$  is assumed. For  $A_1 = 1$ , this gives the maximum number of discretization points as  $M_{max,1} = 50$ .

Applying the optimal  $TS(1)$  sampling design procedure, yields the optimal  $TS(1)$  sampling as shown in Figure 5.6 with  $T^* = 2.27s$  and  $M_1^* = 39$ . With this design aliasing of the assumed compact spectral support is avoided as shown in Figure 5.7(a).

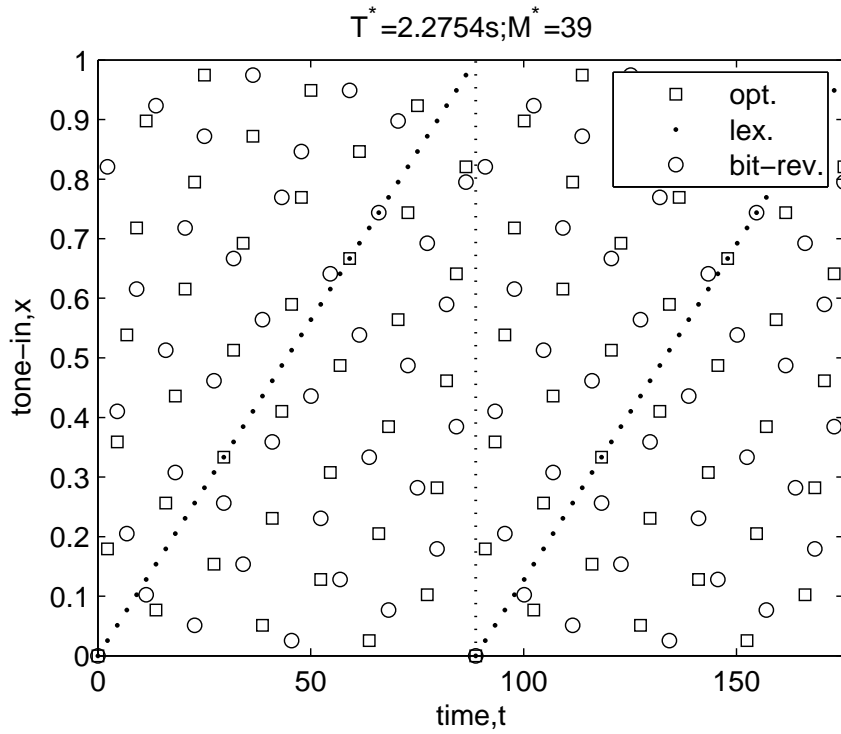


Figure 5.6: The  $M_1^* T^*$  periodic ( $\square$ ) optimal  $TS(1)$  sampling, designed with  $W = 0.04$  [Hz],  $U_1 = 3.0$  [cycles/toner];  $A_1 = 1$ ;  $M_{max,1} = 50$  and  $n = 1$ . Also shown are the  $M_1^* T^*$  periodic ( $\circ$ ) pseudo bit-reversed  $TS(1)$ , and ( $\cdot$ ) lexicographic  $TS(1)$  samplings

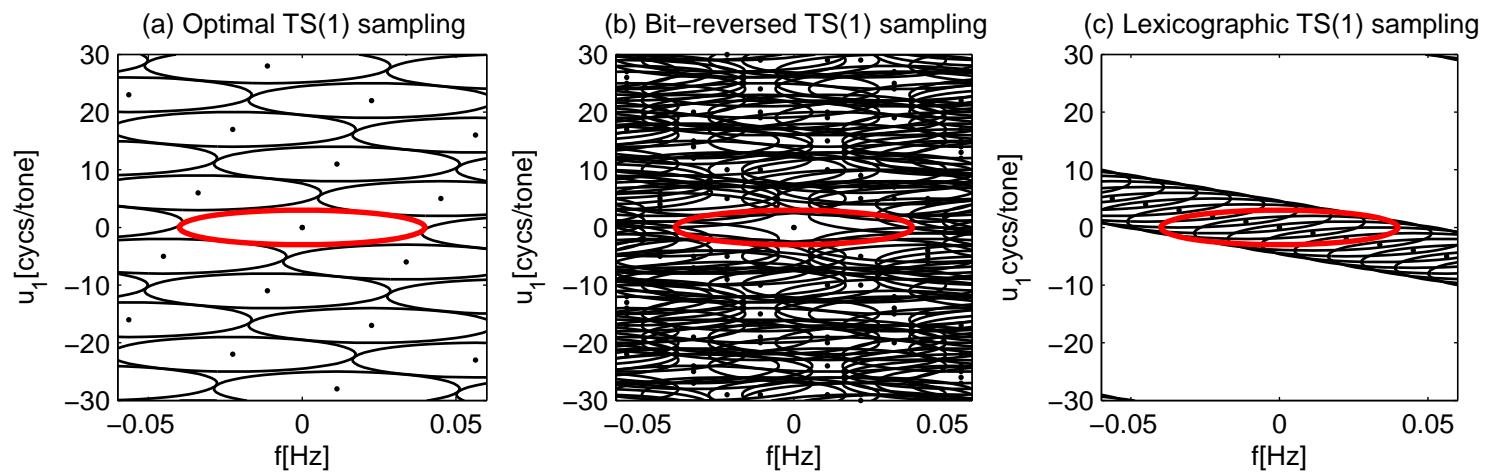


Figure 5.7:  $\Omega_s(u_1, f)$  with (a) optimal  $TS(1)$  sampling; (b) pseudo bit-reversed  $TS(1)$  sampling (note that each of the frequency replication for the pseudo bit-reversed sequence is weighed by some values of  $|Q_{mp}|^2 > 0.01$ ); (c) lexicographic  $TS(1)$  sampling

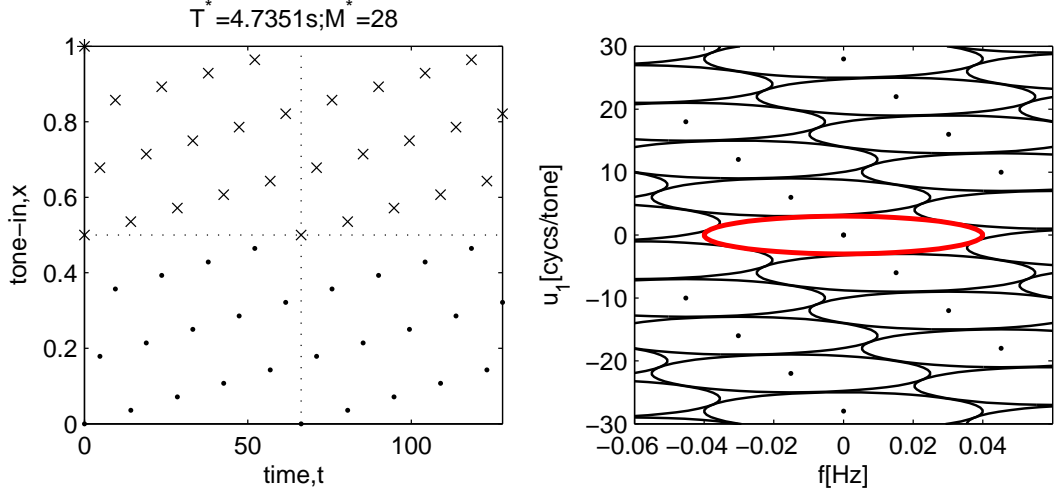


Figure 5.8:  $TS(2)$  sampling sequence (the two samples shown by  $\cdot$ ,  $\times$  at each print cycle) (left) and the corresponding frequency content replication for each signal segment (right)

If the optimal  $TS(1)$  sequence,  $\alpha_1^*(k)$  was replaced by the classical pseudo bit-reversed<sup>4</sup> and lexicographic sequences [11] (see Figure 5.6), aliasing will occur in the temporal-spatial frequency domain as shown in Figure 5.7(b) and (c) respectively. Hence, sampling using the optimal  $TS(1)$  sampling yields superior performance when subsequently reconstructed with a periodic Kalman filter.

If the print cycle time is greater than the 2.27s that is required for the optimal  $TS(1)$  sampling design to work,  $TS(n)$  sampling for  $n > 1$  can be designed. Here, we consider the  $TS(2)$  sampling design. The result of this design is shown in Figure 5.8. The inter-sampling time is approximately two-fold increase compared to the optimal  $TS(1)$  design. This is because by increasing the number of samples at each print cycle, we have more information and therefore we can sample less often. In general, with the  $TS(n)$  sampling design, it is possible to arrive at a compromise between sensors hardware and inter-sampling time requirement.

### 5.6.2 $TS(n)$ sampling design for CRC

Consider the  $TS(n)$  sampling design for the case of the CRC with temporal-spatial dimension  $1 + N = 4$ . Assume in this case that the ellipsoidal temporal-spatial compact spectral support  $\Theta(W, \mathbf{U})$  is parameterized by  $W = 0.01$  [Hz] and  $\mathbf{U} = [2.0, 2.0, 2.0]$  [cycles/tone]. In addition the maximum spatial range,  $\mathbf{A} = [1, 1, 1]$  corresponding to the maximum input

<sup>4</sup>The pseudo bit-reversed sequence is a truncation of the actual periodic bit-reversed sequence given in [11] because  $M_1^*$  is not an exact multiple power of 2

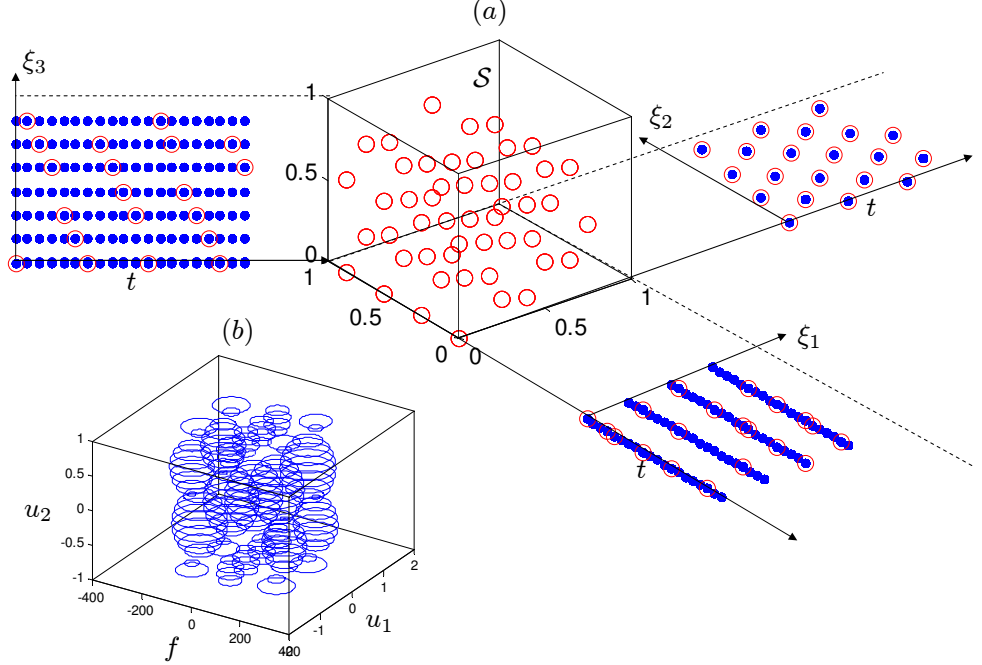


Figure 5.9: (a) The  $\text{prod}(\mathbf{M}^*)T^*$  periodic optimal  $TS(1)$  sampling sequence, designed with  $W = 0.01$  [Hz],  $\mathbf{U} = [2.0, 2.0, 2.0]$  [cycles/tone];  $\mathbf{A} = [1, 1, 1]$ ;  $\mathbf{M}_{max} = [8, 8, 8]$  and  $n = 1$ . (b) The frequency packing projected onto  $(f, u_1, u_2)$  frequency plane at  $u_3 = 0$  which shows no overlapping of the projected frequency content of  $\Theta(W, \mathbf{U})$ .

values of the CMY colorspace. The maximum sensor resolution,  $\varepsilon = [0.125, 0.125, 0.125]$  is assumed. For  $\mathbf{A} = [1, 1, 1]$ , this gives a maximum number of discretization points at each spatial dimension as  $\mathbf{M}_{max} = [8, 8, 8]$ .

Applying the optimal  $TS(1)$  sampling design procedure, yields the optimal  $TS(1)$  sampling as given in Table 5.1. The sequence is shown in Figure 5.9. The design ensure that aliasing of the assume compact spectral support,  $\Theta(W, \mathbf{U}) \subset \mathbb{R} \times \mathbb{R}^3$  is avoided. A  $TS(2)$  sampling can also be designed by taking  $\boldsymbol{\rho} = [2, 1, 1]$ . The results of this design are given in Table 5.1. Similar to the case of the TRC, this design results in more than two-fold increase in the inter-sampling time as compare to the optimal  $TS(1)$  sampling design. Again the desired compromise between sensors hardware and inter-sampling time can be achieved.

Table 5.1: Comparison of different non-aliasing sampling design approaches

$1 + N$	parameters /variables	non-aliasing sampling design			
		opt. $TS(1)$	$TS(2)$	nyquist $TS(1)$	nyquist FULL
$\dagger 2$	$T$	2.28	4.74	2.08	12.50
	$M_1$	39	28	6	6
	$\mathbf{H}_{LTM}$	$\begin{pmatrix} 1 & 0 \\ -7 & 39 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ -5 & 14 \end{pmatrix}$	-	-
	$n$	1	2	1	6
$\ddagger 4$	$T$	1.31	2.71	0.78	50.0
	$[M_1, M_2, M_3]$	[4,7,7]	[8,8,7]	[4,4,4]	[4,4,4]
	$\mathbf{H}_{LTM}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -5 & 0 & 7 & 0 \\ -2 & -7 & -6 & 28 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ -1 & -1 & 2 & 0 \\ -46 & -49 & -14 & 56 \end{pmatrix}$	-	-
	$n$	1	2	1	64

$\dagger$  Design variables:  $W = 0.04$  [Hz],  $U_1 = 3.0$  [cycles/tone],  $A_1 = 1$  and  $M_{max,1} = 50$ . For extension to  $TS(2)$  sampling, set  $\rho_1 = 2$ .

$\ddagger$  Design variables:  $W = 0.01$  [Hz],  $[U_1, U_2, U_3] = [2.0, 2.0, 2.0]$  [cycles/tone],  $[A_1, A_2, A_3] = [1, 1, 1]$  and  $[M_{max,1}, M_{max,2}, M_{max,3}] = [8, 8, 8]$ . For extension to  $TS(2)$  sampling, set  $[\rho_1, \rho_2, \rho_3] = [2, 1, 1]$ .

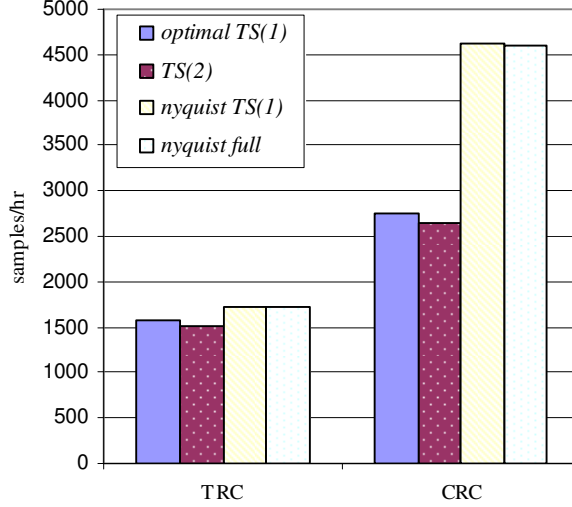


Figure 5.10: Comparison of sensing efficiency

### 5.6.3 Efficiency of the $TS(n)$ sampling design

In this section, the  $TS(n)$  sampling design will be shown to be more effective than equivalent non-aliasing sampling design of the  $1 + N$  dimensional temporal-spatial signal. Efficiency is measured in term of the required samples per unit time. Two non-aliasing sampling designs are considered here by enforcing the temporal-spatial Nyquist conditions:

1. For the case of  $TS(1)$  sampling, by (3.30) pick  $M_i = \lceil A_i U_i \rceil$  for  $i = 1, \dots, N$  and hence, by (3.31) pick  $T = 1/(2WM_t)$ . Notice that by picking these parameters, aliasing will be avoided regardless of sampling sequences,  $\alpha(k)$ .
2. A full sampling approach can also be designed to avoids aliasing by uniformly discretizing the spatial domain with  $M_i = \lceil A_i U_i \rceil$  points for  $i = 1, \dots, N$  and sampling all these discretized samples at each  $T = 1/(2W)$  time interval.

The sampling efficiency comparison between the different methods of sampling derived from the results in Table 5.1 is given in Figure 5.10 for the case of the TRC ( $1 + N = 2$ ) and CRC ( $1 + N = 4$ ).

Both these Nyquist design requires 9.6% and 67.9% more samples per unit time than the optimal  $TS(1)$  design for the case of the TRC and CRC respectively. The optimal  $TS(1)$  sampling achieves a better sampling efficiency by over-sampling the spatial dimension. By trading-off the spatial bandwidth for temporal bandwidth, it is possible to achieve a higher sampling efficiency.



#### 5.6.4 Performance of the $TS(n)$ sampling design

To evaluate the performance of the  $TS(n)$  sampling design, consider the following:

1. The  $N$  dimensional time-dependent vector of uniformly discretized output (each of the  $i$  th spatial dimension is uniformly discretized by  $M_i$  spatial points) of the spatial signal in (3.2) is given by:

$$\bar{\omega}(k) = \bar{\omega}^* + \bar{\mathbf{d}}(k)$$

This form is obtained from (3.4) and (3.7) by taking  $\bar{\mathbf{v}}(k) = \mathbf{0}$ . The disturbance,  $\bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$  where  $M_t = \text{prod}(\mathbf{M})$  is approximated by a discrete-time pink noise dynamics as given by (3.11).

2. A periodic Kalman filter is used to reconstruct the time-sequentially temporal-spatial signal yielding the reconstructed signal,  $\hat{\omega}(k) = \hat{\mathbf{d}}(k) + \bar{\omega}^*$  as given by equations (3.35)-(3.38).
3. To evaluate the reconstruction performance, consider the ratio of norm square reconstruction error to the norm squared deviation from the nominal spatial signal,  $\bar{\omega}^*$  given by:

$$\phi(k) := \frac{\|\bar{\omega}(k) - \hat{\omega}(k)\|_2^2}{\|\bar{\omega}(k) - \bar{\omega}^*\|_2^2} \quad (5.55)$$

The overall reconstruction performance is given by the time-normalized  $\phi(k)$  given by:

$$\Phi_{k \in [l_0, l_0+L]} = \sum_{k \in [l_0, l_0+L]} \phi(k)/L \quad (5.56)$$

For the case of the CRC, it is also meaningful to look at the reconstruction performance in term of the mean color difference formulation of all the  $M_t = \text{prod}(\mathbf{M})$  color values,  $E[\Delta E_{94}^*(k)]$  at each print cycle,  $k$ .

4. The following sparse sampling approaches are considered:
  - (a) Optimal  $TS(1)$  sampling with  $M^*$  points spatial (tonal) discretization and inter-sampling time,  $T^*$ .
  - (b) Lexicographic  $TS(1)$  sampling with  $M^*$  points spatial (tonal) discretization and inter-sampling time,  $T^*$ .
  - (c) Bit-reversed  $TS(1)$  sampling with  $M^*$  points spatial (tonal) discretization and inter-sampling time,  $T^*$ .
  - (d)  $TS(2)$  sampling with  $M^{**}$  points spatial (tonal) discretization and inter-sampling time,  $T^{**}$ .

Table 5.2: Performance of different sampling approaches

$W$ [Hz]	$U_1$ [cycles/tone]	$\Phi_{k \in [l_0, l_0+L]}, l_0 = 500, L = 500 (\times 10^{-3})$					
		opt. $TS(1)$	lex. $TS(1)$	br. $TS(1)$	$TS(2)$	2 fix-pnt samp	3 fix-pnt samp
0.04	3	7.70	31.44	8.51	8.40	34.22	22.32
0.04	4	3.56	14.42	4.08	7.43	28.11	19.97
0.04	5	1.58	4.87	1.77	3.80	14.90	10.32

- (e)  $n$ -fixed point sampling,  $n > 1$  with  $M^{**}$  points spatial (tonal) discretization and inter-sampling time,  $T^{**}$ .

The performance of the  $TS(n), n > 1$  sampling for the case of the TRC (i.e.  $1 + N = 2$  dimensional temporal-spatial signal) is first demonstrated. The following parameters are assumed in the simulation:

- (i) The low-frequency band dynamics of the disturbance is given by (3.12) with a temporal cutoff frequency  $cf_1 = 0.05W$  [Hz] and  $\mathbf{R}_1 = 0.01\mathbf{I}$ .
- (ii) The high-frequency band dynamics of the disturbance is given by (3.13) with temporal  $f$  of a 10-th order low-pass filter at the spatial channel,  $u_1$  given by an ellipse:

$$(u_1/\bar{U}_1)^2 + (f/\bar{W})^2 = 1$$

where  $\bar{U}_1 = 0.8U_1$  [cycles/toner] and  $\bar{W} = 0.8W$  [Hz] gives the highest spatial (tonal) and temporal frequencies, and  $\mathbf{R}_2 = 0.02\mathbf{I}$ . A scaled factor of the actual support size is used here to account for the frequency drop off characteristic of the non-ideal low-pass filter used to approximate the frequency content of time-varying TRC.

Table 5.2 gives the reconstruction performance results for the different sampling approaches and spectral support,  $\Theta(W, \mathbf{U})$  assuming negligible sensing noise. The noise covariance matrix is given by  $\mathbf{R}_{nn} = 1.0 \times 10^{-4}\mathbf{I}$ . The designed optimal  $TS(1)$  and  $TS(2)$  sampling give the overall best performance for the case of  $n = 1$  and  $n > 1$  sample points respectively. It verifies the performance gain in avoiding aliasing of the assumed compact spectral support in the design of the  $TS(n)$  sampling.

We also compared the toner images with the reconstructed TRC using the optimal and lexicographic  $TS(1)$  sampling and a 3-fixed point sampling (REC image) with that of the toner image using fully sampled TRC (FULL image) at a selected print cycle. The 3-fixed point sampling is of great interest here as it is the current method used for the sampling










	3 fix-pnt	opt. TS(1)	lex. TS(1)
REC image			
FULL image			
$\Delta E_{94}^*$ map			

Figure 5.11: The toner image with the reconstructed TRC (REC image) (top); the toner image with the full sampled TRC (FULL image) at a selected print cycle (middle); the  $\Delta E_{94}^*$  map comparing the color tone difference between the REC image and the FULL image (bottom)

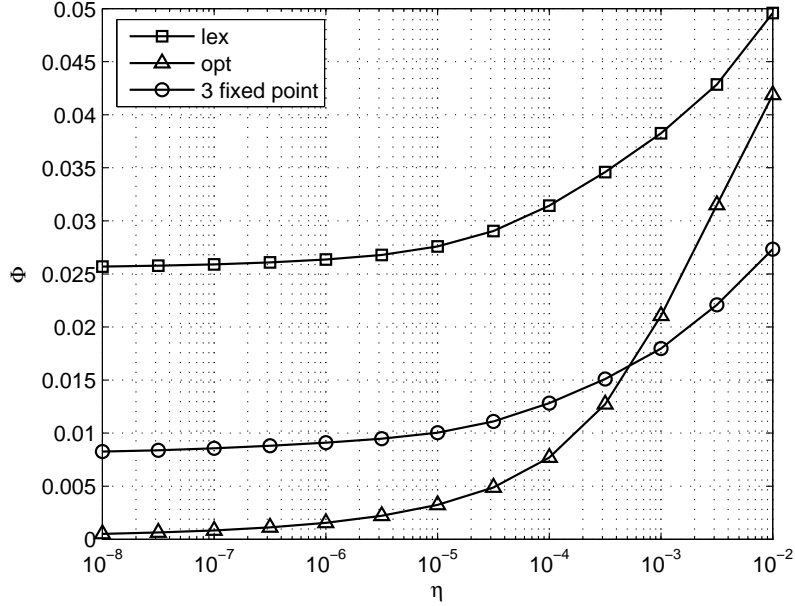


Figure 5.12: The reconstruction performance ratio,  $\Phi_{k \in [l_0, l_0+L]}$ ,  $l_0 = 500, L = 500$  versus different degree of sensing noise,  $\eta$  for different sampling strategies

of the TRC [1]. These images are shown in Figure 5.11. For a quantitative comparison, the  $\Delta E_{94}^*$  (see Appendix A.1) color difference formulation between the REC image and the FULL image is used. Small noticeable difference between the REC image with the optimal  $TS(1)$  sampling and the FULL image is observed. The use of 1 sample at each print cycle gives a reconstructed TRC that is as effective as 39 samples at each print cycle. On the other hand the use of the lexicographic  $TS(1)$  sampling and a 3 fixed point sampling resulted in undesirable perceptually noticeable tone variations compared to the FULL image.

The reconstruction performance depends on the severity of the sensing noise. Figure 5.12 shows the performance ratio,  $\Phi_{k \in [l_0, l_0+L]}$ ,  $l_0 = 500, L = 500$  of the Kalman filter using the 3-fixed point sampling and  $TS(1)$  sampling with optimal and lexicographical sampling sequences at different degree of simulated sensing noise. The severity of the sensing noise is given by  $\eta \in \mathbb{R}$  where the noise covariance matrix,  $\mathbf{R}_{nn_1} = \eta \mathbf{I}$ . For  $\eta \leq 5.2 \times 10^{-4}$ , the optimal  $TS(1)$  sampling gives the best reconstruction performance among the tested sampling strategies. Hence, the optimal  $TS(1)$  sampling makes good use of the uncorrupted information. With increased sensing noise ( $\eta > 5.2 \times 10^{-4}$ ), the 3-fixed point sampling strategy gives a better reconstruction performance. Hence, the ability to recover the time-varying TRC depends not only on how best we make use of the limited information but also the quality of the information itself.

To experimentally evaluate the performance of the  $TS(n)$  sampling, a Xerox Phaser 7700 printer is used to print the color tone patches as specified by the  $TS(n)$  sampling patterns and a X-Rite DTP70 scanning spectrophotometer is then used to measure the color patches in  $L^*a^*b^*$  colorspace. For the case of the TRC, a parameterizations procedure is used to represent the spectrophotometer readings as a one dimensional output tone (see Appendix A.2). Figure 5.13 shows the severity of the sensing noise of the experimental setup for the case of the TRC. Based on this result, the noise covariance matrix is given by  $\mathbf{R}_{nn} = 3.88 \times 10^{-4}\mathbf{I}$ .

Instead of relying on the natural print variations of the printer, the time-varying TRC is artificially induced i.e. it is simulated by disturbance model given by (3.11) with parameters similar to that used in the simulation. This approach is taken because in our existing setup an internal automatic tone correction system is in place and cannot be disabled. Furthermore, by artificially inducing temporal and tonal variations, we can design causal reconstruction filter that accurately reflects the actual temporal-tonal variations. This eliminates the need to consider issue of modeling mismatch, leaving us to concentrate on investigating the performance of the  $TS(n)$  samplings.

The resulting performance of the optimal  $TS(1)$  sampling, lexicographic  $TS(1)$  sampling, the 3 fixed points sampling and full sampling given by  $\Phi_{k \in [l_0, l_0+L]}$ ,  $l_0 = 100, L = 100$  are shown in Figure 5.14. For a fair comparison, for the case of full sampling, a Kalman filter is designed to filter the sensing noise. This case study shows that the reconstructed TRC using the optimal  $TS(1)$  sampling is as good as if we have full measurements. Since one sample tone is needed with the optimal  $TS(1)$  sampling approach and 39 tone samples are needed with the full sampling approach, the optimal  $TS(1)$  sampling is indeed more efficient.

The same experiment is extended for the case of the CRC. To generate the injected disturbances, the following parameters are used:

- (i) The low-frequency band dynamics is given by (3.12) with a temporal cutoff frequency  $cf_1 = 0.1W$  [Hz] and  $\mathbf{R}_1 = 0.001\mathbf{I}$ .
- (ii) The high-frequency band dynamics of the disturbance is given by (3.13) using a bank of 4-th order low-pass filters that filter each tonal frequency channel (each Fourier coefficient) with temporal cutoff frequencies corresponding to an ellipse:

$$\Theta(\bar{W}, \bar{\mathbf{U}}) = \left\{ (f, \mathbf{u}) \left| f^2/\bar{W}^2 + \sum_{i=1}^N u_i^2/\bar{U}_i^2 = 1 \right. \right\}$$

where  $\bar{\mathbf{U}} = 0.8\mathbf{U}$  [cycles/tone] and  $\bar{W} = 0.8W$  [Hz] gives the highest spatial (tonal)

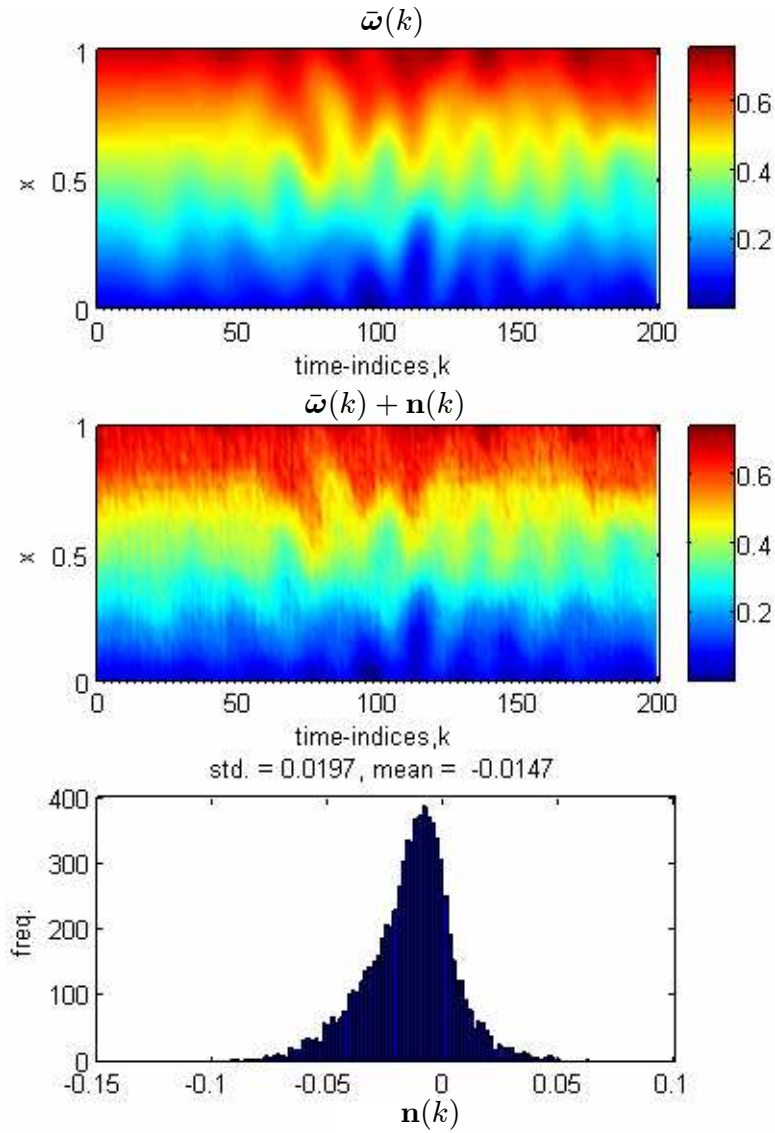


Figure 5.13: The simulated time-varying TRC,  $\bar{\omega}(k) \in \mathbb{R}^{M=39}$  (top figure); the resulting time-varying TRC after printing and sensing,  $\bar{\omega}(k) + \mathbf{n}(k) \in \mathbb{R}^{39}$ , where  $\mathbf{n}(k) \in \mathbb{R}^{39}$  gives the sensing noise (middle figure); the distribution of the sensing noise (bottom figure)

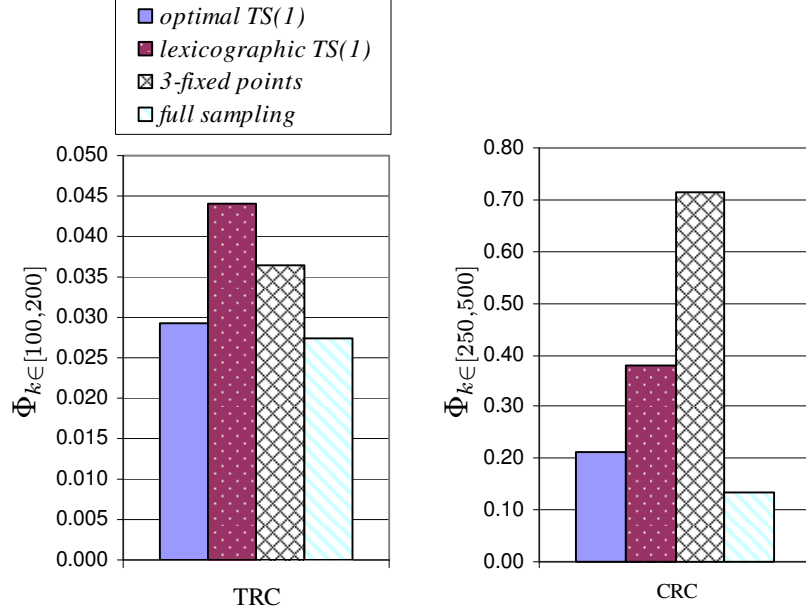


Figure 5.14: Comparison of TRC sensing performance

and temporal frequencies. A scaled factor of the actual support size is used here to account for the frequency drop off characteristic of the non-ideal low-pass filter used to approximate the frequency content of the injected disturbances. Also, the white process noise covariance is given by  $\mathbf{R}_2 = 2.0 \times 10^4 \mathbf{I}$ .

Based on our calibrated sensor setup, the noise characteristics are as shown in Figure 5.15. Based on this result, the noise covariance matrix is given by:

$$\mathbf{R}_{nn} = \begin{pmatrix} \sigma_{L^*}^2 \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_{a^*}^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_{b^*}^2 \mathbf{I} \end{pmatrix} \in \mathbb{R}^{nM_t \times nM_t}$$

where  $\sigma_{L^*}$ ,  $\sigma_{a^*}$  and  $\sigma_{b^*}$  are the standard deviation of sensing noise on the  $L^*$ ,  $a^*$  and  $b^*$  coordinates respectively. Figure 5.16 gives the reconstruction performance comparing the optimal  $TS(1)$  sampling with the lexicographic sequence using  $\phi(k)$  (see equation (5.55)) and the average of color difference formulation,  $E[\Delta E_{94}^*(k)]$  at each print cycle,  $k \in \mathbb{Z}^+$ . The resulting performance of the optimal  $TS(1)$  sampling, lexicographic  $TS(1)$  sampling, the 3 fixed points sampling and full sampling given by  $\Phi_{k \in [l_0, l_0+L]}$ ,  $l_0 = 250$ ,  $L = 250$  are shown in Figure 5.14. For a fair comparison, for the case of full sampling, a Kalman filter is designed to filter the sensing noise. Clearly as in the case of the TRC, the performance of the designed optimal  $TS(1)$  sampling is much better than that of the lexicographic  $TS(1)$

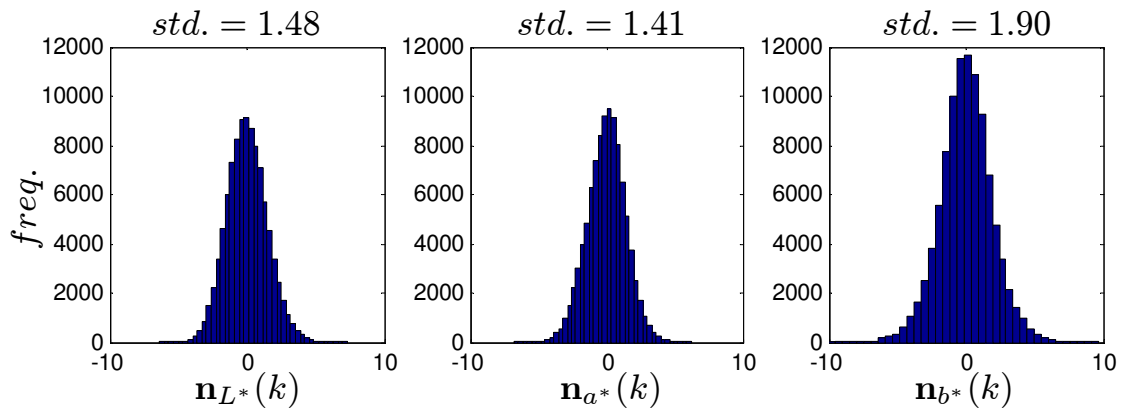


Figure 5.15: The distribution of sensing noise of the  $L^*$ ,  $a^*$  and  $b^*$  color output of the time-varying CRC

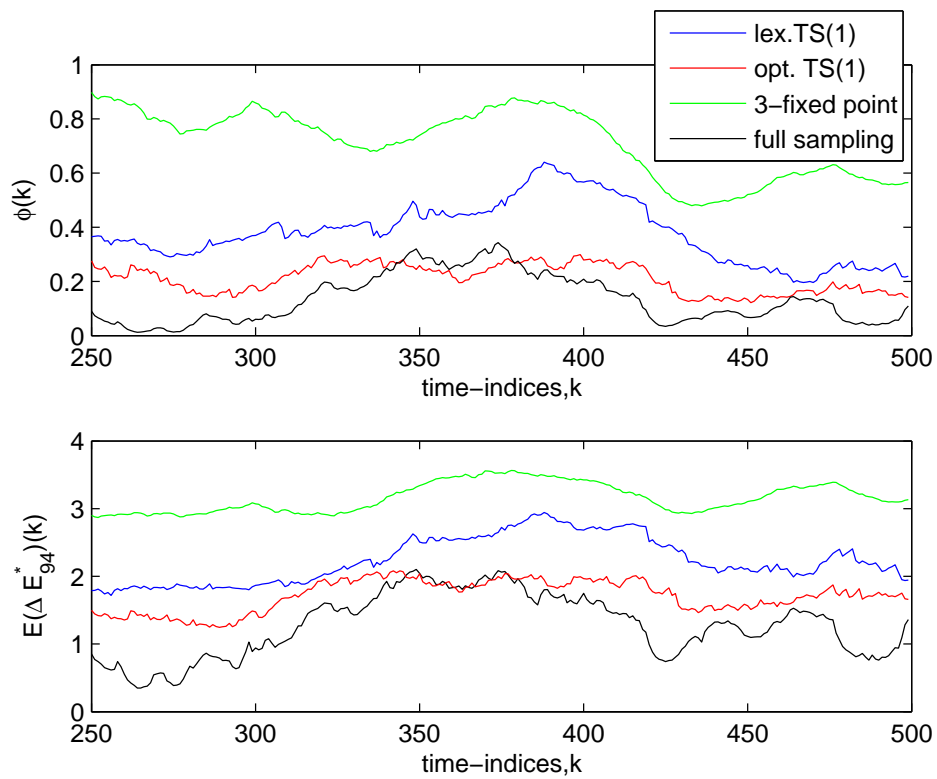


Figure 5.16: Comparing the performance of using the optimal  $TS(1)$  sampling, lexicographic  $TS(1)$  sampling, 3-fixed point sampling and full sampling in sensing the time-varying CRC



sampling and the 3 fixed point sampling and is as good as if we have full sampling of the CRC.

## 5.7 Summary

A design process to obtain the optimal  $TS(1)$  sampling is proposed based on lattice theoretic framework for multidimensional signal processing. Subject to the assumption that all the  $TS(1)$  sampling points lies in a rectangular grid and that the maximum resolution of the spatial grid is upper bounded, it is shown here that the space of all feasible lattices for optimization is compact. Hence a solution to the optimal  $TS(1)$  sampling design that maximizes the inter-sampling time  $T$  and at the same time minimized aliasing in the sampled signal frequency domain is made possible. The designed optimal  $TS(1)$  sampling approach gives a complete description of time-sequential sampling and yield performance and efficiency that is better than a comparable design based on pseudo bit-reversed and lexicographic sequences. The proposed design approach can also be extended for higher spatial dimension and used in the design of  $TS(n)$  sampling for any  $n$  spatial samples at each time instance (print cycle for the case of the TRC and CRC). Therefore the versatility of the proposed approach would be useful as an active sensing strategy that finds applicability in many other spatially distributed systems.

## Chapter 6

# Color Consistency Control

The goal of the xerographic color consistency control system is to ensure that the color reproduction characteristics (CRC) function,  $CRC : \mathcal{C} \rightarrow \mathcal{C}$ , desired-color  $\mapsto$  printed-color is as close to the identity map as possible at all times. From (2.4), the CRC map is the composition of several process steps:

$$CRC(t) = \underbrace{f_{percept}}_{\text{human}} \circ \underbrace{f_{comb} \circ f_{mark}}_{\text{printer}} \circ \underbrace{f_{htone} \circ f_{sep}}_{\text{image processing}}$$

Hence to have the CRC map close to identity at all times, we control the map induced by the xerographic process ( $f_{comb} \circ f_{mark}$ ) via *full* closed loop feedback and we change the image processing processes ( $f_{htone} \circ f_{sep}$ ) via feed-forward such that on nominal values of the human visual perception model ( $f_{percept}$ ), the CRC is close to identity. The proposed color control strategy is shown in Figure 6.1. Unlike the control objective for most processes which is to control or regulate the output of the process, the color control problem consists of maintaining the *process itself* to be constant and stable. The difference is because every customer image to be printed can contain *many* and *any* possible colors which the xerographic printer must reproduce correctly all at once. Moreover, xerographic printers are often used in an on-demand manner in which consecutive customer images are different. Based on the understanding of the underlying digital color printing process as discussed in Chapter 2, we can maintain and stabilize the color consistency of the xerographic digital color printing process by fulfilling the following control objectives:

1. Match the TRC of each  $j$ -th separations to the desired nominal TRC,  $TRC_j^*$  at each

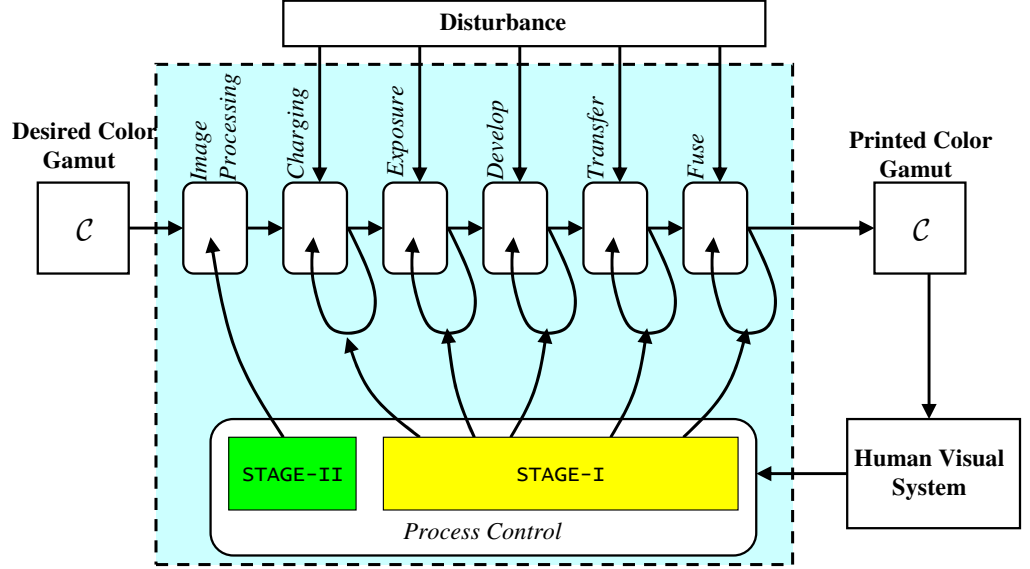


Figure 6.1: Proposed color consistency control strategy

tone,  $x_i \in \mathcal{T}, i = 1, 2, \dots, M_j$ , as  $k \rightarrow \infty$

$$TRC_j(k)[x_i] \rightarrow TRC_j^*[x_i] \quad (6.1)$$

If the manner in which the primary colors are combined is stable and constant, then the output color will also be consistent when the TRC for each primary color separation has been effectively stabilized.

2. Match the desired nominal CRC,  $CRC^*$  at each input colors,  $\mathbf{x}_i \in \mathcal{C}, i = 1, 2, \dots, M_1 M_2 M_3$ , as  $k \rightarrow \infty$

$$CRC(k)[\mathbf{x}_i] \rightarrow CRC^*[\mathbf{x}_i] \quad (6.2)$$

While both control objectives are proposed to achieve color consistency, stabilizing the CRC directly has the added advantage of compensating for the color mixing induced variations. Color mixing is a complex process that is dictated by possibility of mis-registration of the different primary layers and disturbances in the color fusing process (typically through heating). Therefore, variations of colors can occur despite having all the primary TRCs stabilized. Hence, direct CRC control that enables full feedback color control system will likely be more effective.

Both these control objectives are realized in this dissertation with a two-stage control strategy as shown in Figure 6.1. An inner control loop (**STAGE-I**) utilizes the xerographic actuators to regulate the nominal print behavior. Let  $m_j$  denotes the number of xerographic

actuators at each  $j$ -th primary color separation. Since there are fewer xerographic actuators, than the number of tones (i.e.  $m_j \ll M_j$  for each  $j$ -th primary color separation) and colors (i.e.  $m_1 + m_2 + m_3 \ll M_1 M_2 M_3$  for the case of the CRC) to be controlled, it is typically not possible for (6.1) and (6.2) to hold for all the tones and colors in the presence of disturbances. One possibility is to require that (6.1) is satisfied only at  $m_j$  instead of all  $M_j$  tones. Similarly we require that (6.2) is satisfied only at  $m_1 + m_2 + m_3$  instead of all  $M_1 M_2 M_3$  colors. Theoretically, this can be achieved using integral control for constant or slowly varying disturbances. However, as shown in [3] for the case of the TRC, the integral control approach may lead to mis-behavior at the unspecified tones and poses robustness problem in the presence of model uncertainty. Instead, in this dissertation a curve-fitting approach is proposed to minimize the two-norm error of the TRC/CRC over the entire tone/color range.

While we expect that **STAGE-I** control significantly improves the tone consistency subject to the presence of disturbances, the small number of xerographic actuators do not completely eliminate the TRC/CRC deviations from the desired TRC/CRC. By preserving the TRC/CRC range with **STAGE-I** control (in addition to maintaining the TRC/CRC to be relatively stable), we can introduce an outer control loop (**STAGE-II**) that makes use of the image processing process to compensate for the residual TRC/CRC variations.

In both **STAGE-I** and **STAGE-II** control, full feedback information (i.e. the time-varying TRC and CRC map) is available through the sensing approach presented in Chapter 3, 4 and 5. In section 6.1, the **STAGE-I** color stabilization control strategy using a small collection of xerographic actuators to compensate for the disturbances on the xerographic printing process is proposed. An optimal with integral dynamics control and robust static control approaches are presented. Section 6.2 proposed the color feed-forward control strategy (**STAGE-II**) to compensate for the residual color variations through the image processing processes. In section 6.3 the performance of the proposed TRC/CRC stabilization control system is demonstrated. Finally, some concluding remarks are given in section 6.4.

## 6.1 STAGE-I Color Stabilization Control Strategy

In this section, we proposed **STAGE-I** control approach where the  $m_t$  number of xerographic actuators ( $m_t = m_j$  for the  $j$ -th TRC and  $m_t = m_1 + m_2 + m_3$  for the case of the CRC) are used to compensate for the print variations due to the disturbances. From (3.7), both the TRC and CRC can be modeled by the linear static uncertain model:

$$\bar{\mathbf{e}}(k) = \hat{\boldsymbol{\phi}}(\mathbf{I} + \boldsymbol{\Delta}(k)\mathbf{W}_v)\bar{\mathbf{v}}(k) + \bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t} \quad (6.3)$$

where  $M_t = M_j$  for the  $j$ -th TRC with  $N = 1$  (corresponding to a single output tone dimension) and  $M_t = M_1 M_2 M_3$  with  $N = 3$  (corresponding to the three dimension output colorspaces e.g. L\*a\*b\*) for the case of the CRC. In addition  $\bar{\mathbf{v}}(k) \in \mathbb{R}^{m_t}$  is the xerographic actuators input, and  $\bar{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$  is a slowly time varying disturbances.  $\hat{\boldsymbol{\phi}} \in \mathbb{R}^{NM_t \times m_t}$  is the nominal sensitivity function,  $\boldsymbol{\Delta}(k) \in \mathbb{R}^{m_t \times m_t}$  is the multiplicative uncertainty,  $\mathbf{W}_v \in \mathbb{R}^{m_t \times m_t}$  is the matrix of given uncertainty weights. The high dimensionality of the TRC/CRC coupled with limited actuation does not permit us to control all  $M_t$  tones/colors. Even when each CMY color coordinate is modestly discretized into 16 steps, the color quality of  $16^3 = 4096$  desired colors need to be kept track of for color control problem.

In section 6.1.1 and 6.1.2, an optimal with integral dynamics control and a robust static control approaches respectively, are proposed to ensure the  $M_t$ -tones/colors are close to the nominal TRC/CRC respectively, in a least-squared sense. Section 6.1.3 determines the weighting matrix to specify the relative importance of the CRC error at different colors.

### 6.1.1 Linear quadratic with integrator dynamics controller

In this section, the linear quadratic with integrator dynamic color consistency control is presented. Assuming that the plant uncertainty is negligible i.e.  $\boldsymbol{\Delta}(k) = \mathbf{0}$  in (6.3), this control approach allows  $q_t < m_t$  tones/colors to be precisely controlled by imposing integrator dynamics on those  $q_t$  tones/colors on the optimal control formulation. This feature is useful when it is desirable to have certain  $q_t$  tones/colors to converge exactly to the desired tones. However this comes at the expense of reducing the freedom to minimize the error two-norm between the actual TRC/CRC and the desired TRC/CRC. Hence the optimal control problem is to find the control  $\bar{\mathbf{v}}(k)$  based on the measured TRC/CRC,  $\bar{\boldsymbol{\omega}}(k)$ , such that the following quadratic performance index (QPI),  $J$  is minimized:

$$J = \frac{1}{2} \sum_{k \in \mathbb{Z}} \bar{\mathbf{s}}^T(k) \mathbf{W}_s \bar{\mathbf{s}}(k) + \frac{1}{2} \sum_{k \in \mathbb{Z}} \bar{\mathbf{e}}^T(k) \mathbf{W}_e \bar{\mathbf{e}}(k) \quad (6.4)$$

where  $\bar{\mathbf{s}}(k) \in \mathbb{R}^{q_t}$  is the integrated error at the  $q_t < m_t$  tones/colors to be controlled precisely, whose dynamics are given by:

$$\bar{\mathbf{s}}(k+1) = \bar{\mathbf{s}}(k) + \mathbf{C}_s \bar{\mathbf{e}}(k) \quad (6.5)$$

where  $\mathbf{C}_s \in \mathbb{R}^{q_t \times NM_t}$  is the indicator matrix for the selected  $q_t$ -tones/colors to fulfill (6.1),(6.2) respectively.  $\mathbf{W}_s \in \mathbb{R}^{q_t \times q_t}$  and  $\mathbf{W}_e \in \mathbb{R}^{NM_t \times NM_t}$  are the weighting matrices.

**Controller 6.1:** *Linear Quadratic with Integrator Dynamics*

Assuming the disturbance  $\bar{\mathbf{d}}(k)$  is available (it will be replaced by its estimate using the time-sequential sampling and reconstruction approach proposed in the earlier chapters), the optimal control is given by:

$$\bar{\mathbf{v}}(k) = -\mathbf{K}_1 \bar{\mathbf{s}}(k) + \mathbf{K}_2 \bar{\mathbf{d}}(k) \quad (6.6)$$

where

$$\mathbf{K}_1 = \mathbf{Z}_{ww}^{-1} \mathbf{Z}_{xw}^T \quad (6.7)$$

$$\mathbf{K}_2 = -\mathbf{Z}_{ww}^{-1} (\hat{\phi}^T \mathbf{C}_s^T \mathbf{S}_B \mathbf{C}_s + \hat{\phi}^T \mathbf{W}_e - \hat{\phi}^T \mathbf{C}_s^T \mathbf{F}_B) \quad (6.8)$$

are the feedback and feed-forward gains respectively, with

$$\begin{aligned} \mathbf{Z}_{ww} &:= \hat{\phi}^T \mathbf{W}_e \hat{\phi} + \hat{\phi}^T \mathbf{C}_s^T \mathbf{S}_B \mathbf{C}_s \hat{\phi} \\ \mathbf{Z}_{xw} &:= \mathbf{S}_B \mathbf{C}_s \hat{\phi} \end{aligned}$$

$\mathbf{S}_B$  is obtained from the solution of the discrete algebraic Riccati equation.

$$\mathbf{I}^T \mathbf{S}_B \mathbf{I} - \mathbf{S}_B - \mathbf{S}_B \mathbf{C}_s \hat{\phi} \cdot (\hat{\phi}^T \mathbf{W}_e \hat{\phi} + \hat{\phi}^T \mathbf{C}_s^T \mathbf{S}_B \mathbf{C}_s \hat{\phi})^{-1} (\mathbf{S}_B \mathbf{C}_s \hat{\phi})^T + \mathbf{W}_s = \mathbf{0} \quad (6.9)$$

and,

$$\mathbf{F}_B = (\mathbf{Z}_{xw} \mathbf{Z}_{ww}^{-1} \hat{\phi}^T \mathbf{C}_s^T)^{-1} \cdot (\mathbf{Z}_{xw} \mathbf{Z}_{ww}^{-1} (\hat{\phi}^T \mathbf{C}_s^T \mathbf{S}_B \mathbf{C}_s + \hat{\phi}^T \mathbf{W}_e) - \mathbf{S}_B \mathbf{C}_s) \quad (6.10)$$

**Remark 6.1** *To implement (6.6), the Kalman filter estimate of the disturbance signal based on time-sequential sampling,  $\hat{\mathbf{d}}(k) \in \mathbb{R}^{NM_t}$  in (3.35), is used in lieu of the actual disturbance  $\bar{\mathbf{d}}(k)$ .  $\bar{\mathbf{s}}(k)$  is obtained from (6.5) where the Kalman estimate of the temporal-spatial signal error based on time-sequential sampling,  $\hat{\mathbf{e}}(k) \in \mathbb{R}^{NM_t}$  in (3.38), is used in lieu of the actual error,  $\bar{\mathbf{e}}(k)$ .*

**Derivation:**

The linear-quadratic state feedback follower-controller for the given QPI for system (6.3) for  $\Delta(k) = \mathbf{0}$  can be solved by using the backward-sweep solution [39]. Derivation of the

optimal control for TRC/CRC stabilization is as follows. From (6.5):

$$\bar{\mathbf{s}}(k+1) = \bar{\mathbf{s}}(k) + \mathbf{C}_s \bar{\mathbf{e}}(k)$$

For  $\Delta(k) = \mathbf{0}$ , from (6.3) we have:

$$\bar{\mathbf{e}}(k) = \hat{\phi} \bar{\mathbf{v}}(k) + \bar{\mathbf{d}}(k) \quad (6.11)$$

The QPI is given by:

$$J = \frac{1}{2} \bar{\mathbf{s}}^T(K) \mathbf{W}_s \bar{\mathbf{s}}(K) + \frac{1}{2} \sum_{k=0}^{K-1} [\bar{\mathbf{s}}^T(k) \mathbf{W}_s \bar{\mathbf{s}}(k) + \bar{\mathbf{e}}^T(k) \mathbf{W}_e \bar{\mathbf{e}}(k)] \quad (6.12)$$

The integrator dynamics (6.5) and plant (6.11) are adjoint to the quadratic performance index (6.12) with a Lagrange multiplier vector sequence  $\boldsymbol{\lambda}^T(k+1)$  as follows:

$$\bar{J} = J + \sum_{k=0}^{K-1} \boldsymbol{\lambda}^T(k+1) [\bar{\mathbf{s}}(k) + \mathbf{C}_s \hat{\phi} \bar{\mathbf{v}}(k) + \mathbf{C}_s \bar{\mathbf{d}}(k) - \bar{\mathbf{s}}(k+1)] + \boldsymbol{\lambda}^T(0) [\bar{\mathbf{s}}_0 - \bar{\mathbf{s}}(0)]$$

Define the discrete Hamiltonian

$$\mathbf{H}(k) := \frac{1}{2} \left[ \begin{array}{l} \bar{\mathbf{s}}^T(k) \mathbf{W}_s \bar{\mathbf{s}}(k) + \bar{\mathbf{v}}^T(k) \hat{\phi}^T \mathbf{W}_e \hat{\phi} \bar{\mathbf{v}}(k) + \bar{\mathbf{v}}^T(k) \hat{\phi}^T \mathbf{W}_e \bar{\mathbf{d}}(k) \\ + \bar{\mathbf{d}}(k) \mathbf{W}_e \hat{\phi} \bar{\mathbf{v}}(k) + \bar{\mathbf{d}}(k) \mathbf{W}_e \bar{\mathbf{d}}(k) \end{array} \right] + \boldsymbol{\lambda}^T(k+1) [\bar{\mathbf{s}}(k) + \mathbf{C}_s \hat{\phi} \bar{\mathbf{v}}(k) + \mathbf{C}_s \bar{\mathbf{d}}(k)]$$

then:

$$\bar{J} = \frac{1}{2} \bar{\mathbf{s}}^T(K) \mathbf{W}_s \bar{\mathbf{s}}(K) + \sum_{k=0}^{K-1} [\mathbf{H}(k) - \boldsymbol{\lambda}^T(k) \bar{\mathbf{s}}(k)] - \boldsymbol{\lambda}^T(K) \bar{\mathbf{s}}(K) + \boldsymbol{\lambda}^T(0) \bar{\mathbf{s}}_0 \quad (6.13)$$

Consider infinitesimal changes  $\delta \bar{J}$  due to infinitesimal changes in  $\delta(\bar{\mathbf{s}}(k))$ ,  $\delta(\bar{\mathbf{s}}(K))$ ,  $\delta(\bar{\mathbf{s}}_0)$ ,  $\delta(\bar{\mathbf{v}}(k))$  and  $\delta(\bar{\mathbf{d}}(k))$  away from their optimal values:

$$\begin{aligned} \delta \bar{J} = & \bar{\mathbf{s}}^T(K) \mathbf{W}_s \delta(\bar{\mathbf{s}}(K)) - \boldsymbol{\lambda}^T(K) \delta(\bar{\mathbf{s}}(K)) + \boldsymbol{\lambda}^T(0) \delta(\bar{\mathbf{s}}_0) \\ & + \sum_{k=0}^{K-1} [(\mathbf{H}_{\bar{\mathbf{s}}}(k) - \boldsymbol{\lambda}^T(k)) \delta(\bar{\mathbf{s}}(k)) + \mathbf{H}_{\bar{\mathbf{v}}}(k) \delta(\bar{\mathbf{v}}(k)) + \mathbf{H}_{\bar{\mathbf{d}}}(k) \delta(\bar{\mathbf{d}}(k))] \end{aligned} \quad (6.14)$$

The necessary conditions are given by:

1.  $\boldsymbol{\lambda}^T(k) \equiv \mathbf{H}_{\bar{\mathbf{s}}}(k)$ , then:

$$\boldsymbol{\lambda}(k) \equiv \boldsymbol{\lambda}(k+1) + \mathbf{W}_s \bar{\mathbf{s}}(k) \quad (6.15)$$

where  $\boldsymbol{\lambda}^T(K) = \mathbf{W}_s \bar{\mathbf{s}}(K)$

2. For a given initial condition  $\bar{\mathbf{s}}_0$ ,  $\delta(\bar{\mathbf{s}}_0) = 0$ .

3.  $\mathbf{H}_{\bar{\mathbf{v}}}(k) \equiv 0$  and  $\mathbf{H}_{\bar{\mathbf{d}}}(k) \equiv 0$ , then:

$$-\hat{\boldsymbol{\phi}}^T \mathbf{W}_e \hat{\boldsymbol{\phi}} \bar{\mathbf{v}}(k) = \hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \boldsymbol{\lambda}(k+1) + \hat{\boldsymbol{\phi}}^T \mathbf{W}_e \bar{\mathbf{d}}(k) \quad (6.16)$$

Consider the backward sweep solution of the following form:

$$\boldsymbol{\lambda}(k) = \mathbf{S}_B(k) \bar{\mathbf{s}}(k) - \boldsymbol{\lambda}_B(k) \quad (6.17)$$

From condition (6.15) we have:

$$\boldsymbol{\lambda}(k) = (\mathbf{S}_B(k+1) + \mathbf{W}_s) \bar{\mathbf{s}}(k) + \mathbf{Z}_{xw} \bar{\mathbf{v}}(k) - \boldsymbol{\lambda}_B(k+1) + \mathbf{S}_B(k+1) \mathbf{C}_s \bar{\mathbf{d}}(k) \quad (6.18)$$

From condition (6.16) we have:

$$0 = \mathbf{Z}_{xw}^T \bar{\mathbf{s}}(k) + \mathbf{Z}_{ww} \bar{\mathbf{v}}(k) - \hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \boldsymbol{\lambda}_B(k+1) + [\hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \mathbf{S}_B(k+1) \mathbf{C}_s + \hat{\boldsymbol{\phi}}^T \mathbf{W}_e] \bar{\mathbf{d}}(k) \quad (6.19)$$

where  $\mathbf{Z}_{ww} := \hat{\boldsymbol{\phi}}^T \mathbf{W}_e \hat{\boldsymbol{\phi}} + \hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \mathbf{S}_B(k+1) \mathbf{C}_s \hat{\boldsymbol{\phi}}$ ;  $\mathbf{Z}_{xw} := \mathbf{S}_B(k+1) \mathbf{C}_s \hat{\boldsymbol{\phi}}$ . Substituting (6.19) into (6.18) and comparing it with the form given by (6.17) gives:

$$\mathbf{S}_B(k) = \mathbf{S}_B(k+1) + \mathbf{W}_s - \mathbf{Z}_{xw} \mathbf{Z}_{ww}^{-1} \mathbf{Z}_{xw}^T \quad (6.20)$$

where  $\mathbf{S}_B(K) = \mathbf{W}_s$

$$\begin{aligned} \boldsymbol{\lambda}_B(k) = & [\mathbf{I} - \mathbf{Z}_{xw} \mathbf{Z}_{ww}^{-1} \hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T] \boldsymbol{\lambda}_B(k+1) + \\ & [\mathbf{Z}_{xw} \mathbf{Z}_{ww}^{-1} (\hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \mathbf{S}_B(k+1) \mathbf{C}_s + \hat{\boldsymbol{\phi}}^T \mathbf{W}_e) - \mathbf{S}_B(k+1) \mathbf{C}_s] \bar{\mathbf{d}}(k) \end{aligned} \quad (6.21)$$

where  $\boldsymbol{\lambda}_B(K) = 0$

At steady state, we obtained the discrete algebraic Riccati equation from (6.20) as given by (6.9). The solution of the discrete algebraic Riccati equation gives  $\mathbf{S}_B$  and from (6.21) the optimal control is derived:

$$\bar{\mathbf{v}}(k) = -\mathbf{Z}_{ww}^{-1} \mathbf{Z}_{xw}^T \bar{\mathbf{s}}(k) - \mathbf{Z}_{ww}^{-1} (\hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \mathbf{S}_B \mathbf{C}_s + \hat{\boldsymbol{\phi}}^T \mathbf{W}_e - \hat{\boldsymbol{\phi}}^T \mathbf{C}_s^T \mathbf{F}_B) \bar{\mathbf{d}}(k)$$



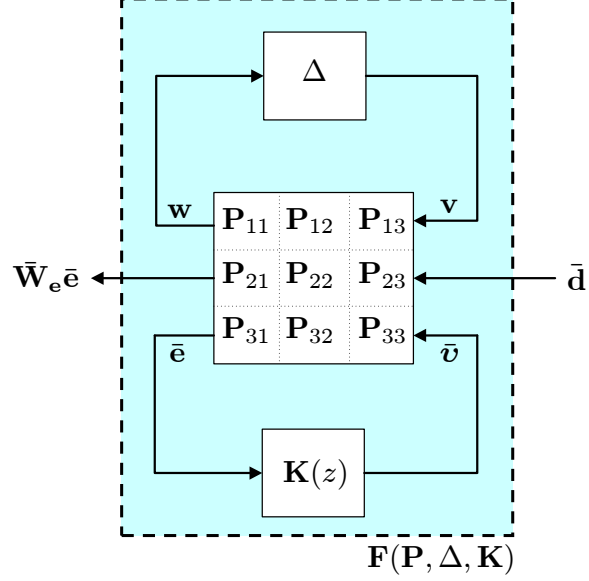


Figure 6.2: LFT representation of linear system model

where  $\mathbf{F}_B$  is given by (6.10), yielding the controller form as given by (6.6).

### 6.1.2 Robust static controller

In this section the robust stabilization control approach for TRC/CRC stabilization proposed in [3] is presented. Consider the linear uncertain model given in (6.3), with  $\Delta(k)$  and  $\bar{\mathbf{d}}(k)$  being effectively unknown. Let  $\bar{\mathbf{U}}(z) = \mathbf{K}(z)\bar{\mathbf{E}}(z)$ <sup>1</sup> be some linear feedback to be specified. Defining the error weighting  $\bar{\mathbf{W}}_e \in \mathbb{R}^{NM_t \times NM_t}$  to specify the relative importance of the TRC/CRC error at different tones/colors. The closed-loop system can be expressed as a linear fractional transformation (LFT) as in Figure 6.2 where  $\mathbf{P}$  is a known matrix. Specifically we have:

$$\begin{pmatrix} \mathbf{w} \\ \bar{\mathbf{W}}_e \bar{\mathbf{e}} \\ \bar{\mathbf{e}} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} & \bar{\mathbf{W}}_v \\ \bar{\mathbf{W}}_e \hat{\phi} & \bar{\mathbf{W}}_e & \bar{\mathbf{W}}_e \hat{\phi} \\ \hat{\phi} & \mathbf{I} & \hat{\phi} \end{pmatrix}}_{\mathbf{P}} \begin{pmatrix} \mathbf{v} \\ \bar{\mathbf{d}} \\ \bar{\mathbf{v}} \end{pmatrix} \quad (6.22)$$

with feedback connections

$$\mathbf{v}(k) = \Delta(k)\mathbf{w}(k)$$

$$\mathbf{U}(z) = \mathbf{K}(z)\bar{\mathbf{E}}(z)$$

<sup>1</sup> $z$  is the  $z$ -transform variable,  $\bar{\mathbf{U}}(z)$  and  $\bar{\mathbf{E}}(z)$  are the  $z$ -transform of the sequences  $\bar{\mathbf{v}}(k)$  and  $\bar{\mathbf{e}}(k)$  respectively

Due to the static nature of the xerographic printing process, and the disturbances are generally slowly varying, the performance optimization is restricted to the steady-state case. Hence, since  $\bar{\mathbf{W}}_e \bar{\mathbf{e}}^\infty$  is linear with respect to  $\bar{\mathbf{d}}_e^\infty$ , there exists some matrix,  $\mathbf{F}(\mathbf{P}, \Delta^\infty, \mathbf{K}^\infty)$  such that:

$$\bar{\mathbf{W}}_e \bar{\mathbf{e}}^\infty = \mathbf{F}(\mathbf{P}, \Delta^\infty, \mathbf{K}^\infty) \bar{\mathbf{d}}_e^\infty$$

The goal here is therefore to find a controller  $\mathbf{K}(z)$  such that for the worst case performance and for as large a class of uncertainty  $\Delta(\cdot)$  as possible  $\|\mathbf{F}(\mathbf{P}, \Delta^\infty, \mathbf{K}^\infty)\|_2$  is minimized. This is achieved by optimizing the following steady-state performance index:

$$\bar{\gamma}(\mathbf{K}^\infty) = \min \left\{ \gamma \left| \sup_{\|\Delta^\infty\| \leq \frac{1}{\gamma}} \bar{\sigma}(\mathbf{F}(\mathbf{P}, \Delta^\infty, \mathbf{K}^\infty)) \leq \gamma \right. \right\}$$

where  $\bar{\sigma}[\cdot]$  denotes the maximum singular value of its argument. Hence, the optimal controller dc gain is

$$\mathbf{K}_\Delta := \arg \min_{\mathbf{K}^\infty} \bar{\gamma}(\mathbf{K}^\infty)$$

Finding  $\mathbf{K}_\Delta$  can be achieved by procedure described in [3, 56].

**Controller 6.2:** *Robust Static Controller Implementation*[3]

The robust static controller is given by:

$$\bar{\mathbf{v}}(k) = \mathbf{K}_\Delta \bar{\mathbf{e}}(k) \tag{6.23}$$

This controller is not realizable since  $\bar{\mathbf{e}}(k)$  is not available until  $\bar{\mathbf{v}}(k)$  is given. Instead, let the realization of the controller be of the form:

$$\bar{\mathbf{v}}(k+1) = \mathbf{A}_\Delta \bar{\mathbf{v}}(k) + \mathbf{B}_\Delta \bar{\mathbf{e}}(k) \tag{6.24}$$

Hence the controller will have a sub-optimal dc gain if

$$\mathbf{K}_\Delta = (\mathbf{I} - \mathbf{A}_\Delta)^{-1} \mathbf{B}_\Delta \tag{6.25}$$

and the nominal stability of the closed-loop system (i.e.  $\Delta(k) = \mathbf{0}$ ) is given by:

$$|\text{eig}(\mathbf{A}_\Delta + \mathbf{B}_\Delta \hat{\phi})| < 1 \tag{6.26}$$

By conditions (6.25) and (6.26), we have  $\mathbf{A}_c = \mathbf{A}_\Delta + \mathbf{B}_\Delta \hat{\phi} = \mathbf{L} + \mathbf{A}_\Delta (\mathbf{I} - \mathbf{L})$  where  $\mathbf{L} = \mathbf{K}_\Delta \hat{\phi}$  is the nominal loop gain. Hence, given that  $\mathbf{A}_c$  is a stable closed-loop matrix (picked by

the designer), we can find the  $\mathbf{A}_\Delta$  and  $\mathbf{B}_\Delta$  matrices from the dynamic controller as follows:

$$\mathbf{A}_\Delta = (\mathbf{A}_c - \mathbf{L})(\mathbf{I} - \mathbf{L})^{-1} \quad (6.27)$$

$$\mathbf{B}_\Delta = (\mathbf{I} - \mathbf{A}_\Delta)\mathbf{K}_\Delta \quad (6.28)$$

**Remark 6.2** *To implement (6.24), the periodic Kalman filter estimate of the temporal-spatial signal error based on TS(n) sampling,  $\hat{\mathbf{e}}(k) \in \mathbb{R}^{NM_t}$  in (3.38), is used in lieu of the actual error,  $\bar{\mathbf{e}}(k)$ . Further details of this controller can be found in [3]. A major issue addressed in [3] is that fixed TRC sampling is assumed, and the number of tones that are measured is small compared to the dimensionality of the TRC ( $n \ll (M_t = M_1)$ ). In this dissertation, this work is extended in two important ways:*

1. *Since only small numbers of fixed tones are used, the TRC is not fully characterized by actual measurements. Hence, the performance at the unmeasured tones is only as good as the model used to describe them. With the use of TS(n) sampling, the apparent number of measured tones/colors can be significantly increased ( $\rightarrow M_t$ ), thus relaxing this difficulty.*
2. *The robust static controller is extended for the stabilization of the CRC map.*

### 6.1.3 Determination of the error weighting matrix

The error weighting matrix given in the linear quadratic formulation,  $\mathbf{W}_e$  (see equation (6.4)) and the robust static formulation,  $\bar{\mathbf{W}}_e$  (see equation (6.22)) is related by:

$$\mathbf{W}_e = \bar{\mathbf{W}}_e^T \bar{\mathbf{W}}_e$$

In this section, the error weighting,  $\mathbf{W}_e \in \mathbb{R}^{3M_t \times 3M_t}$  which specifies the relative importance of the CRC error at different colors is deduced. The error weighting matrix is made up of the color and image specific weighting matrices [32] i.e.:

$$\mathbf{W}_e = \mathbf{W}_{\Delta E^*} \mathbf{W}_\varphi \quad (6.29)$$

where

1.  $\mathbf{W}_{\Delta E^*} \in \mathbb{R}^{3M_t \times 3M_t}$  is a diagonal matrix which gives the color specific weighting. It accounts for the sensitivity of the human visual system (HVS) to colors variations.

2.  $\mathbf{W}_\varphi \in \mathbb{R}^{3M_t \times 3M_t}$  is also a diagonal matrix with diagonal element of either 1 or 0 - 1 indicating nominal colors that are significant in an image,  $\varphi$  and 0 indicating non-significant colors.

### 6.1.3.1 Color specific weighting, $\mathbf{W}_{\Delta E^*}$

To incorporate the sensitivity of the HVS to color variations, it is convenient to specify the output colorspace of the CRC mapping by the CIE L\*a\*b\* colorspace. Let  $L_i^* a_i^* b_i^*(k)$  denotes the  $i$ -th values of the the time-varying output color. In this context, the time dependent error vector of discretized output of the spatial signal is given by:

$$\bar{\mathbf{e}}(k) = [\Delta L_1^*(k), \dots, \Delta L_{M_t}^*(k), \Delta a_1^*(k), \dots, \Delta a_{M_t}^*(k), \Delta b_1^*(k), \dots, \Delta b_{M_t}^*(k)]^T \in \mathbb{R}^{3M_t} \quad (6.30)$$

where  $\Delta L_i^*(k) := L_i^*(k) - L_i^{**}$ ,  $\Delta a_i^*(k) := a_i^*(k) - a_i^{**}$ ,  $\Delta b_i^*(k) := b_i^*(k) - b_i^{**}$ .  $L_i^{**} a_i^{**} b_i^{**}$  gives the  $i$ -th value of the nominal color gamut. In the color-science literature [18], the color difference formulation,  $\Delta E_i^*$  that quantify the difference (distance measure) between  $i$ -th index value of the printed color gamut and the corresponding  $i$ -th index value of the nominal color gamut, can be used to specify the error metric function given by:

$$J_{\Delta E^*}(k) = \sum_{i=1}^{M_t} \Delta E_i^{*2}(k) = \bar{\mathbf{e}}^T(k) \mathbf{W}_{\Delta E^*} \bar{\mathbf{e}}(k) \quad (6.31)$$

In [18], typical color difference formulations are given by:

$$\Delta E_{ab,i}^*(k) = \left[ \Delta L_i^{*2}(k) + \Delta a_i^{*2}(k) + \Delta b_i^{*2}(k) \right]^{\frac{1}{2}} \quad (6.32)$$

$$\Delta E_{94,i}^*(k) = \left[ \left( \frac{\Delta L_i^*(k)}{k_{L,i} S_{L,i}} \right)^2 + \left( \frac{\Delta C_{ab,i}^*(k)}{k_{C,i} S_{C,i}} \right)^2 + \left( \frac{\Delta H_{ab,i}^*(k)}{k_{H,i} S_{H,i}} \right)^2 \right]^{\frac{1}{2}} \quad (6.33)$$

where

$$\Delta C_{ab,i}^*(k) = (a_i^{*2}(k) + b_i^{*2}(k))^{\frac{1}{2}} - (a_i^{**2} + b_i^{**2})^{\frac{1}{2}} \quad (6.34)$$

$$\Delta H_{ab,i}^*(k) = \left[ \Delta a_i^{*2}(k) + \Delta b_i^{*2}(k) - \Delta C_{ab,i}^{*2}(k) \right]^{\frac{1}{2}} \quad (6.35)$$

$$S_{L,i} = 1 \quad (6.36)$$

$$S_{C,i} = 1 + 0.045 C_{ab,i}^{**} \quad (6.37)$$

$$S_{H,i} = 1 + 0.015 C_{ab,i}^{**} \quad (6.38)$$

$$C_{ab,i}^{**} = (a_i^{**2} + b_i^{**2})^{\frac{1}{2}} \quad (6.39)$$

$k_L, k_C, k_H$  are positive real-valued scaling parameters. As noted in [18],  $k_L = k_C = k_H = 1$  at reference conditions <sup>2</sup>.

In (6.32), the two-norm of the  $L^*a^*b^*$  colorspace is used. Hence assuming that we pick all colors in the discretized color gamut, from (6.30),(6.31) and (6.32), the error metric can be restated as:

$$J_{\Delta E_{ab}^*}(k) = \bar{\mathbf{e}}^T(k) \bar{\mathbf{e}}(k)$$

In this case, we have that  $\mathbf{W}_{\Delta E^*} = \mathbf{I}$ . While this is convenient, it is achieved by assuming that the  $L^*a^*b^*$  colorspace is perceptually uniform. It is well-known [34, 18] that the  $L^*a^*b^*$  colorspace is only “nearly” uniform in that equal Euclidean distance between points in the  $L^*a^*b^*$  colorspace at different regions of the color space does not give exact equal perceptual differences. In this case the color difference formulation in (6.33) gives a better quantification of sensitivity of the HVS across the color gamut. Assuming reference conditions and substituting (6.35) into (6.33) we have:

$$\begin{aligned} \Delta E_{94,i}^{*2}(k) &= \left( \frac{\Delta L_i^*(k)}{S_{L,i}} \right)^2 + \left( \frac{\Delta a_i^*(k)}{S_{H,i}} \right)^2 + \left( \frac{\Delta b_i^*(k)}{S_{H,i}} \right)^2 + \\ &\Delta C_{ab,i}^{*2}(k) \left( \frac{1}{S_{C,i}^2} - \frac{1}{S_{H,i}^2} \right) \end{aligned} \quad (6.40)$$

Since the actual colors to be stabilized are expected to be closed to the nominal, it is assumed here that  $\Delta C_{ab,i}^* \approx (\Delta a_i^{*2} + \Delta b_i^{*2})^{\frac{1}{2}}$ . Hence, from (6.40) we have:

$$\Delta E_{94,i}^{*2}(k) \approx \left( \frac{\Delta L_i^*(k)}{S_{L,i}} \right)^2 + \left( \frac{\Delta a_i^*(k)}{S_{C,i}} \right)^2 + \left( \frac{\Delta b_i^*(k)}{S_{C,i}} \right)^2$$

Therefore, the performance index given in (6.31) assuming that we pick all colors in the discretized gamut, is given by:

$$J_{\Delta E_{94}^*}(k) = \sum_{i=1}^{M_t} \Delta E_{94,i}^{*2}(k) = \bar{\mathbf{e}}^T(k) \mathbf{W}_{\Delta E^*} \bar{\mathbf{e}}(k) \quad (6.41)$$

---

<sup>2</sup>The reference conditions assume that a pair of nontextured specimens in edge contact are placed on a background with  $L^* = 50$ , illuminated by D65 with an illuminance of 1000 lux, and viewed at a distance so that the pair subtends a visual angle of at least  $4^\circ$ . Although not explicitly stated, a perceptibility judgment is assumed

where

$$\mathbf{W}_{\Delta E^*} = \text{diag} \begin{pmatrix} 1/S_{L,1}^2 \\ \vdots \\ 1/S_{L,M_t}^2 \\ 1/S_{C,1}^2 \\ \vdots \\ 1/S_{C,M_t}^2 \\ 1/S_{C,1}^2 \\ \vdots \\ 1/S_{C,M_t}^2 \end{pmatrix} \in \mathbb{R}^{3M_t \times 3M_t} \quad (6.42)$$

### 6.1.3.2 Image specific weighting, $\mathbf{W}_\varphi$

The image specific weighting,  $\mathbf{W}_\varphi$  is achieved by a two-step procedure:

1. Let  $\mathcal{C}_\varphi \subset \mathbb{R}^3$  denote the color gamut of the image,  $\varphi$ . In this step, we determine  $\mathcal{C}_\varphi^* \subset \mathcal{C}_\varphi$  such that  $\mathcal{C}_\varphi^*$  gives the collection of colors of image,  $\varphi$  that are visually most significant to a human viewer.
2. Deduce the image specific weighting,  $\mathbf{W}_\varphi$  from the significant color image gamut,  $\mathcal{C}_\varphi^*$ . In general if a collection of  $p$  images (i.e.  $\varphi_1, \dots, \varphi_p$ ) needs to be printed, we deduce  $\mathbf{W}_\varphi$  from  $\bigcup_{i=1}^p \mathcal{C}_{\varphi_i}^*$ .

To determine the visually significant color gamut,  $\mathcal{C}_\varphi^* \subset \mathcal{C}_\varphi$ , consider the visual redundancies on a color image. For a color image, the redundancies can be found in the color and spatial representation of the image. Human visual system (HVS) has lower sensitivity to high frequency components (detail) of high chrominance image. Moreover, typically “smooth” images are visually more irrelevant compared to textured and detailed images. These effects in the color and spatial domain are taken into account to reduce the redundancy in the image representation by following similar procedure to that used in wavelet transform coding [57, 58]. The multi-resolution property of wavelet decomposition is particularly well suited to the properties of the HVS because psycho-visual studies show that the human visual system can be modeled using multiple perceptual channels, which are octave-wise spread over the spatial frequency range. From one octave to the next, the summation area (area of the retina contributing to the perceived signal) also changes by a factor of four. This behavior perfectly matched that of the dyadic structure of the discrete wavelet transform (DWT).

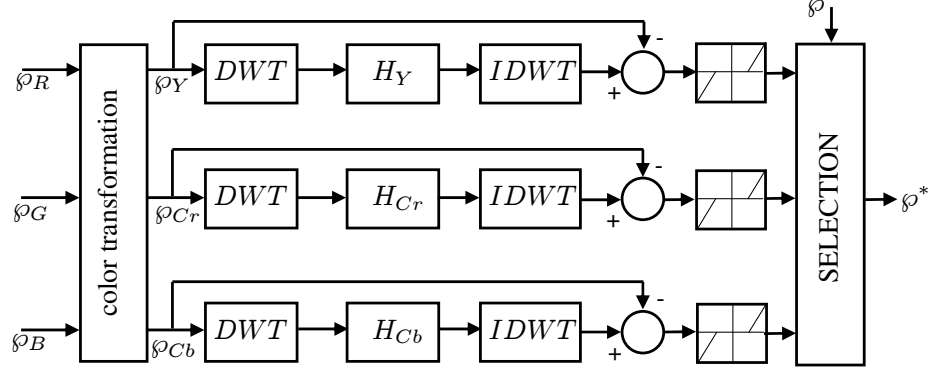


Figure 6.3: Finding the significant color map,  $\varphi^*$ . Here *DWT* denotes the discrete wavelet transform and *IDWT* denotes the inverse discrete wavelet transform

The basic idea of the proposed approach to determine  $\mathcal{C}_\varphi^*$  is shown in Figure 6.3. First a RGB image ( $(\varphi_R, \varphi_G, \varphi_B)$  image plane) is transformed to the YCrCb format ( $\rightarrow (\varphi_Y, \varphi_{Cr}, \varphi_{Cb})$  image plane). This representation is used because the Y, Cr and Cb color planes by themselves are widely de-correlated [34], and hence provide a suitable platform for the development of the algorithm to follow (see Appendix A.1 for RGB $\rightarrow$ YCrCb transformation). The discrete wavelet transform (DWT) with *embedded HVS model* is then performed on each of the image plane individually. The wavelets coefficients of this decomposition are then used to synthesize a color “enhanced” image with the inverse discrete wavelet transform (IDWT).

By comparing the original image at each color plane with that of the reconstructed image, we find areas where there are significant differences. These areas indicate the locations where there are significant colors or spatial information on that color plane. By picking these areas on the original image,  $\varphi$  we have the significant color image,  $\varphi^*$  where  $\mathcal{C}_\varphi^*$  can be obtained.

To embed the HVS into the 2D wavelet decomposition, the contrast sensitivity function (CSF) is used. The CSF quantitatively describes the human eye’s sensitivity to changes in luminance and chrominance as a function of spatial frequency. The shapes of the chrominance CSF’s are different from that of the luminance CSF’s. The following luminance and chrominance CSF’s model due to Marcus [59] is used:

*Luminance CSF:*

$$H_Y(f_s) = 0.997f_s^2 e^{-0.970f_s^{0.758}} + 0.22e^{-0.800f_s^{1.999}} \quad (6.43)$$

*Chrominance CSFs:*

$$H_{Cr}(f_s) = e^{-0.1521f_s^{0.893}} \quad (6.44)$$

$$H_{Cb}(f_s) = e^{-0.2041f_s^{0.900}} \quad (6.45)$$

where  $f_s$  is spatial frequency in cycles per degree [cpd]. The frequency representation in cpd makes the CSF independent of the viewing distance. The CSFs for each Y, Cr and Cb color channels are given in Figure 6.4(a). To make use of these CSFs, it is necessary to know the maximum frequency,  $f_{s_{max}}$  that the human eye can tolerate at a certain viewing distance,  $v$  [m] for a given image resolution,  $r$  [pixel/inch], as follows:

$$f_{s_{max}} = v \tan(0.5^\circ) r / 0.0254$$

Hence, the CSF function that is to be embedded into the wavelet decomposition is dependent upon the viewing distance and the quality of image to be viewed. Typically images with  $r = 72$  [pixel/inch] resolution are viewed from a distance of  $v = 0.3$  [m].

The relationship between the CSF and a 4-level 2D wavelet decomposition is given in Figure 6.4(b). The 2D wavelet decomposition gives the corresponding horizontal and vertical details of the image at each level of decomposition. The frequency range of the sub-band corresponds to the horizontal and vertical frequency range of the CSF function. As an illustration, the shaded area of the CSF function at the vertical and horizontal frequency axis correspond to the LH2 of the wavelet decomposition. The weighting that must be used for the wavelet coefficients at each sub-band is described by the corresponding horizontal and vertical frequency range in the CSF functions. The typical 1D wavelet transform can be extended to a 2D wavelet transform using separable wavelet filters. With separable filters, the 2D wavelet decomposition can be computed by applying a 1D transform to all the rows of the input, and repeating this on all of the columns.

The CSF's embedding with this decomposition is also shown in Figure 6.4(c). The embedded CSF model in this decomposition tree, at each decomposition level,  $\ell$  for the low pass sub-band denoted by  ${}^\ell F_C(\omega)$  and high pass sub-band denoted by  ${}^\ell F_D(\omega)$  are given as follows (see [59] for details):

$${}^\ell F_C(\omega) = H_{cc} \left( \frac{|\omega| f_{s_{max}}}{\pi 2^\ell} \right) \quad (6.46)$$

$${}^\ell F_D(\omega) = H_{cc} \left( \frac{f_{s_{max}}}{2^{\ell-1}} - \frac{|\omega| f_{s_{max}}}{\pi 2^\ell} \right) \quad (6.47)$$



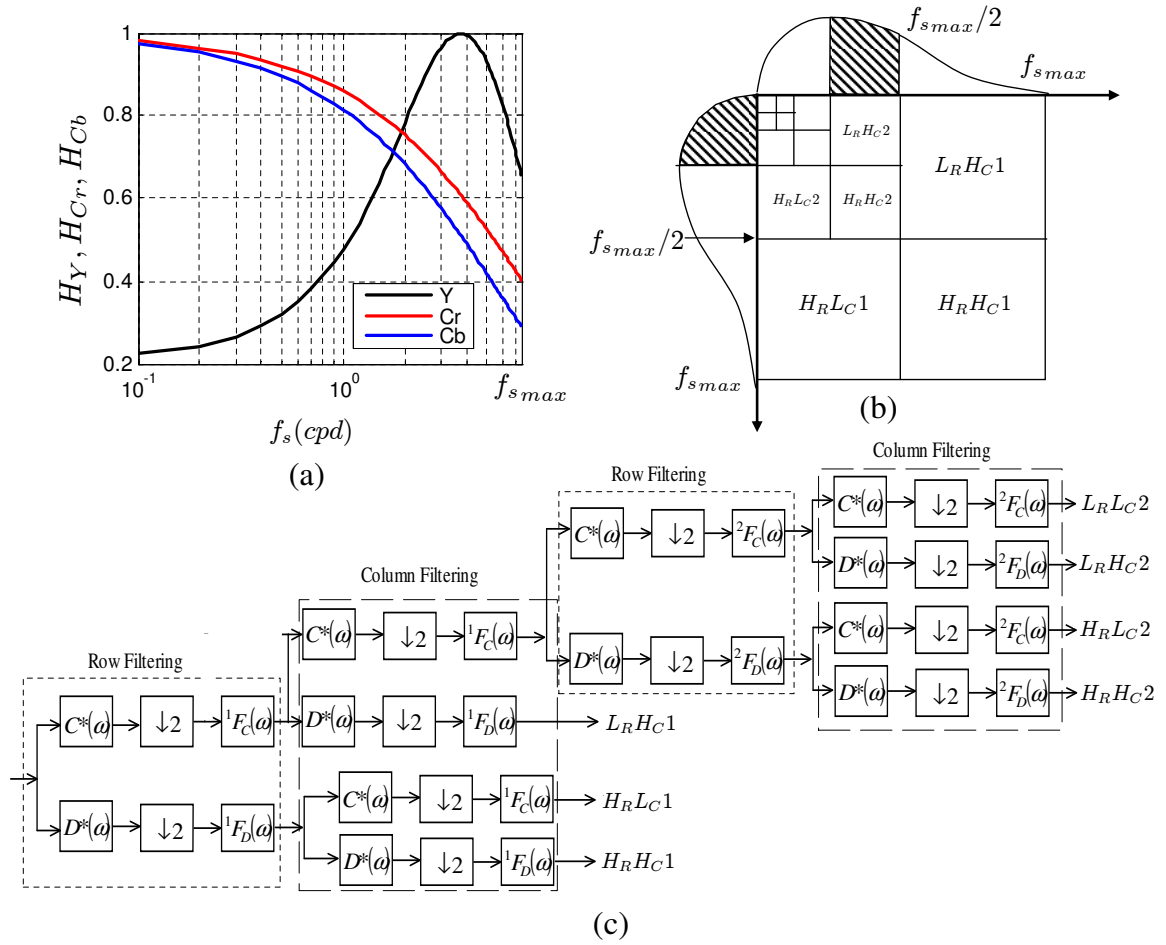


Figure 6.4: (a) The CSF for luminance (Y) and chrominance (Cr,Cb) channels; (b) Relation between the CSF and a 4-level 2D wavelet decomposition. As an illustration, the shaded areas of the CSF function at the vertical and horizontal frequency axis correspond to the  $L_R H_C 2$  of the wavelet decomposition. The weighting that must be used for the wavelet coefficients at each sub-band is described by the corresponding horizontal and vertical frequency range in the CSF functions; (c) 2-level 2D wavelet decomposition with embedded human visual system.

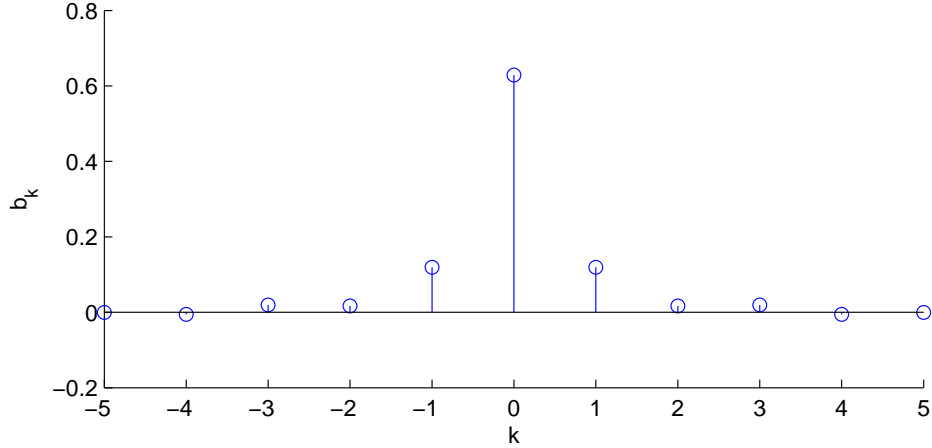


Figure 6.5: The HVS FIR filter coefficients for the low-pass sub-band range at  $\ell = 1$

where  $\omega \in [-\pi, \pi]$  and  $cc = \{Y, Cr, Cb\}$  denotes the color channels with  $H_{cc}$  given by equations (6.43)-(6.45). The embedding of these HVS filters makes the wavelet decomposition as shown in Figure 6.4(c) different from that of standard wavelet decomposition. These filters appropriately weights the spatial frequency band of the image that are critical to a human viewer in each of the de-correlated Y, Cr and Cb color planes.

To implement  ${}^\ell F_C(\omega), {}^\ell F_D(\omega)$  filters and embed them in the wavelet decomposition, they are approximated here by FIR-filters. Note that  ${}^\ell F_C(\omega)$  and  ${}^\ell F_D(\omega)$  have even symmetry (i.e.  ${}^\ell F_C(\omega) = {}^\ell F_C(-\omega)$  and  ${}^\ell F_D(\omega) = {}^\ell F_D(-\omega)$ ), hence they can be expressed by real and even Fourier series:

$${}^\ell F_C(\omega) = {}^\ell b_{C0} + 2 \sum_{k=1}^{n_F} {}^\ell b_{Ck} \cos(k\omega) \quad (6.48)$$

$${}^\ell F_D(\omega) = {}^\ell b_{D0} + 2 \sum_{k=1}^{n_F} {}^\ell b_{Dk} \cos(k\omega) \quad (6.49)$$

where

$$\begin{aligned} {}^\ell b_{D0} &= \frac{1}{\pi} \int_0^\pi {}^\ell F_D(\omega) d\omega & {}^\ell b_{Dk} &= \frac{1}{\pi} \int_0^\pi {}^\ell F_D(\omega) \cos(k\omega) d\omega \\ {}^\ell b_{C0} &= \frac{1}{\pi} \int_0^\pi {}^\ell F_C(\omega) d\omega & {}^\ell b_{Ck} &= \frac{1}{\pi} \int_0^\pi {}^\ell F_C(\omega) \cos(k\omega) d\omega \end{aligned}$$

Figure 6.5 shows the  $n_F = 5$  filter coefficients for  ${}^1 F_C(\omega)$  at a viewing distance,  $v = 0.3m$  and for an image with image resolution,  $r = 72$  pixel/inch. The coefficients of the HVS embedded wavelet decomposition are used to synthesize a new visually enhance image. To

locate the area where significant color and spatial information reside, the enhance image is compared to the original image at each color planes. Let  $\wp_{cc}(x, y)$  be the  $(x, y)$  pixel values of the original image, and  $\wp_{cc}^S(x, y)$  be the  $(x, y)$  pixel values for the synthesized image at a particular color plane,  $cc$ . The difference between the two images at pixel  $(x, y)$  is given as follows:

$$\delta_{cc}(x, y) = \left( \frac{\wp_{cc}(x, y) - \bar{w}}{w_{max} - w_{min}} \right) - \left( \frac{\wp_{cc}^S(x, y) - \bar{z}}{z_{max} - z_{min}} \right)$$

where  $\bar{w} = \text{mean}(\wp_{cc})$ ,  $w_{max} = \max(\wp_{cc})$ ,  $w_{min} = \min(\wp_{cc})$ ,  $\bar{z} = \text{mean}(\wp_{cc}^S)$ ,  $w_{max} = \max(\wp_{cc}^S)$ , and  $w_{min} = \min(\wp_{cc}^S)$ . To obtain the binary significant map,  $\mathfrak{S}_{cc}$  at the particular  $cc$  color plane, thresholding is performed.

$$\mathfrak{S}_{cc}(x, y) = \begin{cases} 1 & \text{for } |\delta(x, y)| > \delta_t \\ 0 & \text{otherwise} \end{cases}$$

where  $\mathfrak{S}_{cc}(x, y)$  gives the binary significant map value at  $(x, y)$  pixel and  $\delta_t$  is the threshold setting.

Figure 6.6 shows the results of this algorithm. The original image is shown on the left with the corresponding color gamut of this image in term of the printer's output color gamut at nominal actuator settings. Applying the algorithm yields the binary significant map for each of the Y, Cr and Cb color plane as shown in the middle. The white region gives the region of interest at each color planes. In the luminance( Y) plane, note that the dark regions of the color image have been selected because it has stark contrast to the surrounding brighter region and appears more intense. Also the red color appears intense in the picture, and this has been selected in the red chrominance (Cr) color plane. By picking the region of interest(white regions on the binary significant map of all the color planes) on the original image,  $\wp$  we have the significant color image,  $\wp^*$  as shown on the right of Figure 6.6. The corresponding significant color gamut,  $\mathcal{C}_{\wp}^*$  is also shown. Clearly, the required colors to consider for stabilization are less compared to the colors in the original image.

Figure 6.7 shows the resulting significant color image with different threshold setting,  $\delta_t$ . A higher threshold setting enables restriction to a smaller set of critical colors for stabilization. Appropriate setting of this threshold depends on trade-off between color consistency requirements (i.e. quality of printouts) and capability of the system to compensate for disturbances.

The image specific weighting,  $\mathbf{W}_{\wp} \in \mathbb{R}^{3M_t \times 3M_t}$  can now be deduced from the significant color image,  $\wp^*$  with its color lying in gamut,  $\mathcal{C}_{\wp}^*$ . This is achieved by placing greater emphasis (weights) on error values of the CRC i.e.  $\bar{\mathbf{e}}(k) \in \mathbb{R}^{3M_t}$  in (6.30), where the nominal color

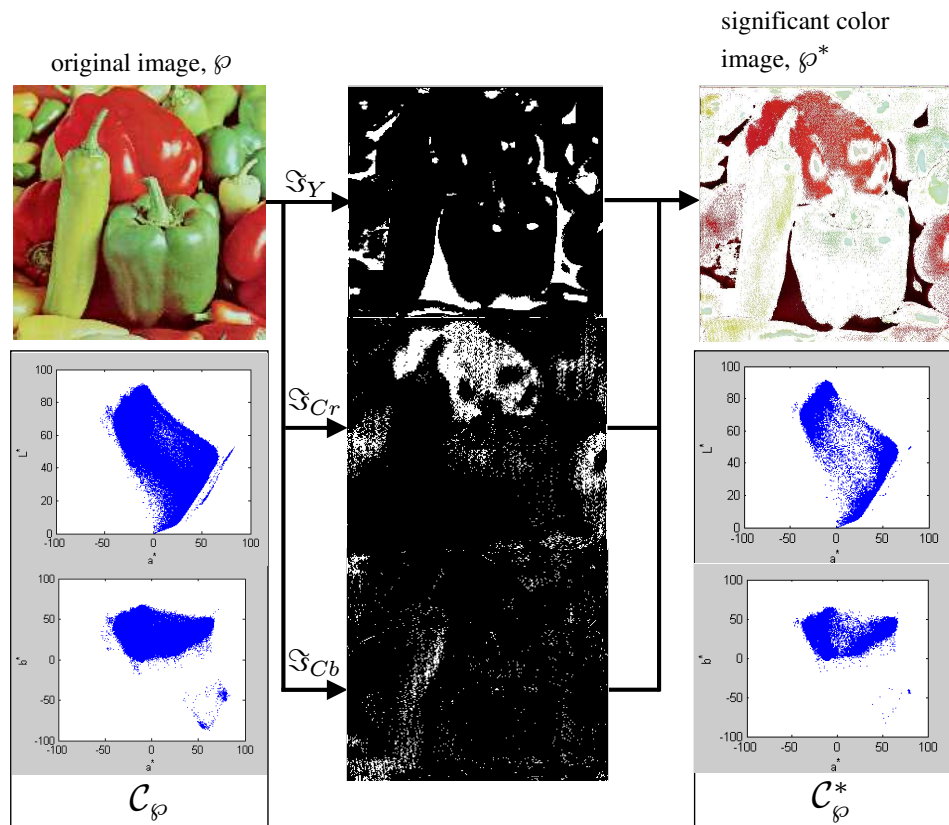


Figure 6.6: The original image,  $\varphi$  with the corresponding desired color gamut in the printer's  $L^*a^*b^*$  colorspace,  $\mathcal{C}_\varphi$  (left); the binary significant map at each YCrCb color planes using  $\delta_t = 3\sigma, \sigma = \text{std}(|\delta_{cc}|)$  (middle); the resulting significant color image,  $\varphi^*$  with the corresponding desired color gamut in the printer's  $L^*a^*b^*$  colorspace,  $\mathcal{C}_\varphi^*$  (right)

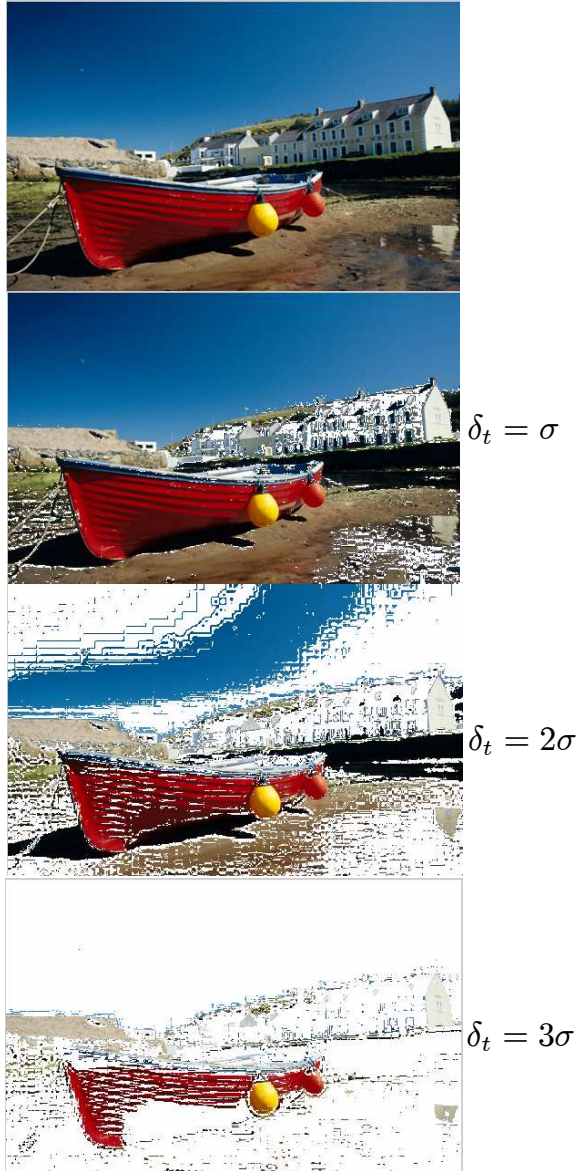


Figure 6.7: Significant color image at different threshold settings

gamut of the print-system is in the significant color gamut of the image,  $\mathcal{C}_\varphi^*$ . Let  $w_\varphi(i)$  be the weight assigned for the  $i$ -th value of the print-system *nominal* color gamut given by  $[L_i^{**}, a_i^{**}, b_i^{**}]$ . Hence, for  $i = 1, \dots, M_t$ ,  $w_\varphi(i)$  is defined by:

$$w_\varphi(i) = \begin{cases} 1 & \text{if } [L_i^{**}, a_i^{**}, b_i^{**}] \in \mathcal{C}_\varphi^* \\ 0 & \text{otherwise} \end{cases}$$

Let  $\mathbf{w}_\varphi = [w_\varphi(1), \dots, w_\varphi(M_t)]$ , then the image specific weighting can now be defined as  $\mathbf{W}_\varphi = \text{diag}(\mathbf{w}_\varphi, \mathbf{w}_\varphi, \mathbf{w}_\varphi)$  corresponding to the weighting of the L\*, a\* and b\* values of the color gamut.

## 6.2 STAGE-II Color Feed-forward Control Strategy

While the STAGE-I color consistency control is expected to significantly improve the color reproduction consistency subject to physical and environmental disturbances, it will not be possible to completely eliminate deviations of the TRC or CRC from the ideal mapping. This is because only a small number of actuators are available. To further improve the color consistency control performance, it is possible to adjust the image processing steps such that the control objectives stated in (6.1) and (6.2) can be fully realized.

Focusing on the case of the CRC (the same idea applied as well for the case of the TRC), from (2.4), the CRC at time  $t \in \mathbb{R}$  is given by:

$$CRC(t) := \underbrace{f_{percept}}_{\text{human}} \circ \underbrace{f_{comb} \circ f_{mark}}_{\text{printer}} \circ \underbrace{f_{htone} \circ f_{sep}}_{\text{image processing}} \quad (6.50)$$

Let  $f_P := f_{percept} \circ f_{comb} \circ f_{mark} \circ f_{htone}$ . In typical xerographic color printer,  $f_P : \mathcal{C}_{\text{CMY}} \rightarrow \mathcal{C}_{L^*a^*b^*}$ . In STAGE-II control we sought to find the inverse print map,  $f_P^{-1} : \mathcal{C}_{L^*a^*b^*} \rightarrow \mathcal{C}_{\text{CMY}}$  such that

$$f_P \circ f_P^{-1} = I$$

This objective is realized by taking  $f_P^{-1}$  as a software profile map (such as the ICC profile [34]) i.e. imposes on the  $f_{sep}$  map in (6.50). This enables the printed colors to correspond to the desired colors. There are two important considerations for this strategy to work.

Firstly, the effectiveness of this strategy is limited to only colors that  $f_P$  (i.e. the printer) can reproduce. Let the color gamut of these reproducible colors be denoted here by  $\widehat{\mathcal{C}}_{L^*a^*b^*} \subset \mathcal{C}_{L^*a^*b^*}$ . To realize effective color consistency control, the extend (or range) of  $\widehat{\mathcal{C}}_{L^*a^*b^*}$  needs to be maximized. This can only be achieved by using the STAGE-I control that directly affects the physical print process. Hence STAGE-I requirement is modified here, so that

in addition to maintaining the  $f_P$  map to be relatively stable, the goal of preserving and maximizing  $\widehat{\mathcal{C}}_{L^*a^*b^*}$  range will also be emphasized. This emphasis can be achieved by appropriate putting higher weights in the weighting matrix,  $\bar{\mathbf{W}}_e$ .

Secondly, to find the inverse print map,  $f_P^{-1}$  at each time,  $t = kT, k \in \mathbb{Z}^+$  consider a set of  $J$  number of desired colors i.e.

$$\mathcal{D} = \left\{ (\mathbf{x}_j, \mathbf{y}_j) \mid \mathbf{x}_j \in \mathcal{C}_{\text{CMY}}, \mathbf{y}_j \in \widehat{\mathcal{C}}_{L^*a^*b^*} \text{ for } j = 1, \dots, J \right\} \quad (6.51)$$

$\mathbf{x}_j$  for  $j = 1, \dots, J$  is assumed to uniformly covers the CMY colorspace. Here we seek to approximate  $f_P^{-1}$  such that

$$\tilde{f}_P^{-1} = \arg \min_{f_P^{-1}} \frac{1}{J} \sum_{j=1}^J \|\mathbf{x}_j - f_P^{-1}(t)[\mathbf{y}_j]\|^2 \quad (6.52)$$

The most common and easiest way to address this problem is to first evaluate the interpolation functions at a regular lattice of points, in the input space (i.e. CMY colorspace) and then build a multidimensional look-up table (LUT) [34]. A fast interpolation technique such as trilinear or tetrahedral interpolation is then used to transform the image data using this LUT, hence realizing the implementation of approximating  $f_P^{-1}$  and imposing it on  $f_{sep}$ . However reasonable accuracy can typically be achieved only at the expense of a large number of training samples (i.e. large  $J$ ). In addition, this method does not directly address the requirement of (6.52). To address these problems, Groff [10] proposed approximation for the forward print map  $f_P$  and its inverse,  $f_P^{-1}$  using a class of piecewise linear functions. A continuously invertible piecewise linear function on a multidimensional space is parameterized by a set of knot points and a connectivity. The knot points are defined by pairs of *vertices* in the domain and codomain spaces. The connectivity defines a triangulation of the domain and codomain, with the knots serving as vertices - triangle for a two dimension space; tetrahedra for a three dimensional space; and, more generally  $N_s$ -simplices for an  $N_s$  dimensional space. For the data set  $\mathcal{D}$  in (6.51), an initial set of knots and specified connectivity, an optimization algorithm (known as *minvar*) seek to improve the initial piecewise linear approximation of the sampled data by iteratively moving the knots such that condition (6.52) is fulfilled. The *minvar* algorithm is a generalization of the one-dimensional graph intersection algorithm.

In this dissertation, we proposed the use of the LUT or piecewise linear approximation method to approximate  $f_P^{-1}$  at each print cycle,  $k$ . The data set  $\mathcal{D}$  in (6.51) needed for this construction is obtained from the reconstructed CRC using the time-sequential sampling

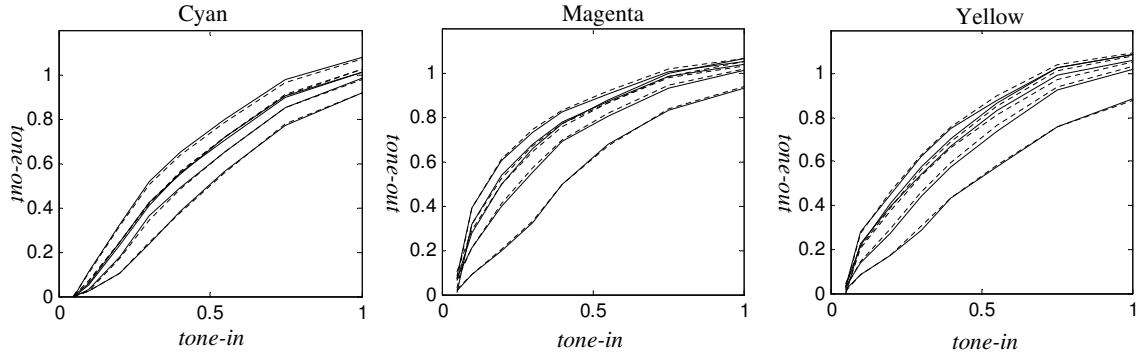


Figure 6.8: Comparison of the actual TRC (solid line) and prediction (dotted line) using the nominal sensitivity at four different sets of actuator settings for the Cyan (left), Magenta (Middle) and Yellow (right) primary color separations. Data for this identification is supplied by Xerox.

approach as proposed in Chapter 3. The use of this reconstructed information avoid the needs to print large number of samples to define the inverse CRC. Additionally, the use of the time-sequential sampling approach enables the right instance to update the inverse print map to ensure that the color reproduction consistency is maintained.

### 6.3 Simulation and Experimental Results

The behavior of the xerographic plant for the case of the TRC is modeled by taking the form given by (6.3) with  $N = 1$ . To identify the matrix of nominal sensitivity function  $\hat{\phi}$ , factorial experiments were performed on the system. In this data, each experiment corresponds to each of the actuators at one of 2 settings within its range. The nominal sensitivity function was then obtained using the least square method for each CMY primary color separation. A comparison of the predicted TRC based on the nominal sensitivity, with the actual TRC at different actuator settings is shown in Figure 6.8. Similarly for the case of the CRC, the behavior of the xerographic color printer is modeled by taking the form given by (6.3) with  $N = 3$ . To identify the matrix of nominal sensitivity function  $\hat{\phi}$ , factorial experiments were performed on the system. In this data each experiment corresponds to each of the actuators of the CMY primary color separations at one of the 2 settings within their range. With  $m = 3$  actuators, we have a total of  $2^3 = 8$  different settings for each of the primaries. With the CMY primary color separations we have a total of  $8^3 = 512$  different set of actuators settings to consider. From these data, the nominal sensitivity function,  $\hat{\phi}$  were then obtained using the least square method.



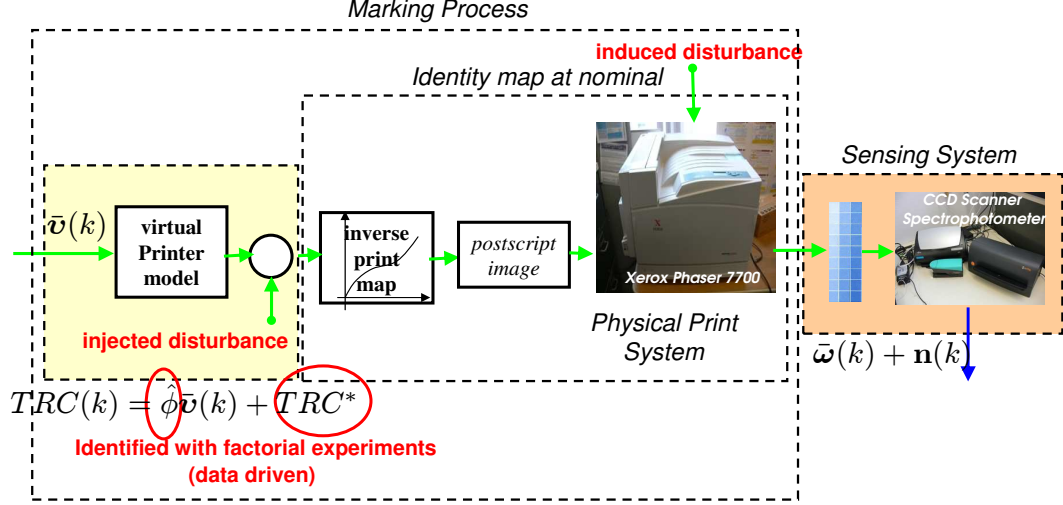


Figure 6.9: Experimental xerographic marking setup for TRC stabilization

The disturbance  $\bar{\mathbf{d}}(k)$  is given in (3.11):

$$\bar{\mathbf{d}}(k) = \mathbf{F}_\ell \cdot \mathbf{x}_\ell(k) + \mathbf{F}_h \cdot \mathbf{x}_h(k) \quad (6.53)$$

where  $k \in \mathbb{Z}^+$  is the time index.  $\mathbf{F}_\ell \cdot \mathbf{x}_\ell(k)$  gives the low-frequency band TRC variations with  $\mathbf{F}_\ell \in \mathbb{R}^{M_t \times M_t}$  denoting the spatial basis function, and  $\mathbf{x}_\ell(k) \in \mathbb{R}^{M_t}$  gives time-varying coefficients modeled by (3.12).  $\mathbf{F}_h \cdot \mathbf{x}_h(k)$  gives the high-frequency band variations with  $\mathbf{F}_h \in \mathbb{R}^{M_t \times M_t}$  denoting the matrix of the Fourier basis functions and  $\mathbf{x}_h(k) \in \mathbb{R}^{M_t}$  is the vector of Fourier coefficients representing the spatial frequency content of the disturbance modeled by a pink noise dynamics as given by (3.13).

### 6.3.1 TRC stabilization control

The proposed TRC stabilization system was experimentally tested on a Xerox Phaser 7700 xerographic printer. A X-rite DTP70 scanning spectrophotometer is used to measure the printed tone patches. Figure 6.9 shows the schematic of our setup. We do not have direct access to the xerographic actuators for this setup. To evaluate the proposed TRC stabilization control system, a virtual printer model is used to generate the response (tonal image) due to changes in the actuator inputs. The virtual printer model is given by (6.3) with no disturbance ( $\bar{\mathbf{d}}(k) = \mathbf{0}$ ) and no plant perturbation ( $\Delta(k) = \mathbf{0}$ ). By calibrating the printer such that it is an identity map at nominal condition, we can capture the effect of the disturbances on the TRC stabilization system. The disturbances were artificially introduced either by introducing a transparency in the optical path of the laser or by injecting

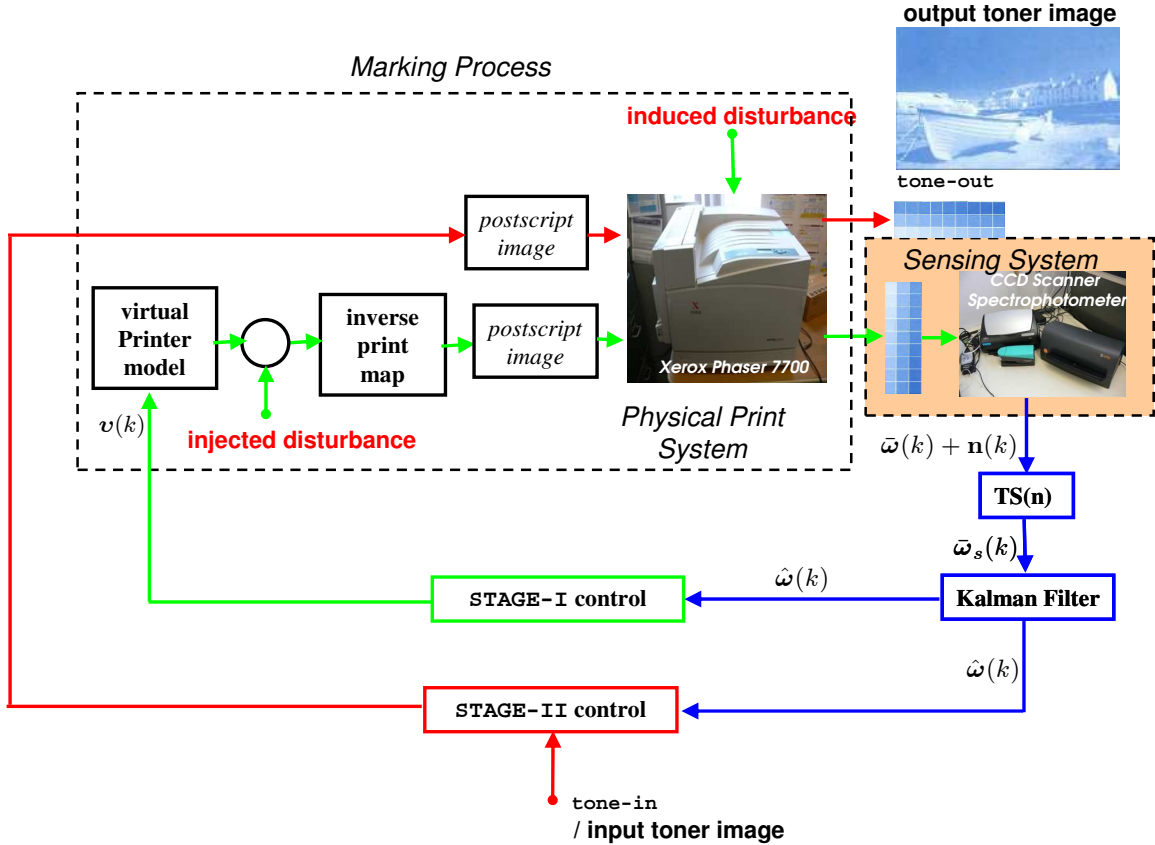


Figure 6.10: Sensing and control for TRC stabilization

a simulated disturbance source  $\bar{\mathbf{d}}_{inject}(k) = \bar{\mathbf{d}}(k)$  (see equation (6.53)) as shown in Figure 6.9.

Figure 6.10 shows the sensing and control setup to experimentally achieve tone consistency. The feedback information is obtained by fully sampling the TRC,  $\bar{\omega}(k) + \mathbf{n}(k)$  or from the reconstructed TRC,  $\hat{\omega}(k)$ . Note that the effectiveness of the the curve-fitting controller (STAGE-I control) depends largely on the degree in which the disturbance signal is in the range space of the sensitivity function,  $\hat{\phi}$ . We assume here that, a relatively good degree of control authority is available to stabilize the TRC. The purpose of the STAGE-I is to maintain the range of the TRC. Residual variations can then be compensated in the image processing actuators with the availability of the reconstructed TRC. The effectiveness of STAGE-II control depends on the accuracy of the reconstructed TRC at all  $M_t$  tones.

### Case Study 1

Simulation and experimental studies were first conducted for STAGE-I control for the case of

the TRC using the lexicographic and bit-reversed  $TS(1)$  sampling approach. The objective of this study is to validate the proposed STAGE-I approach. In this first study, the following specific system behaviors were considered:

- (i) The behavior of the system without plant perturbation is considered i.e.  $\Delta(k) = \mathbf{0}$ .
- (ii) The disturbance  $\bar{\mathbf{d}}(k)$  dynamics is simulated from the high-frequency band pink noise model given in (3.13) with the temporal cutoff frequency,  $f$  [Hz] of the butterworth filter at each spatial channel,  $u_1$  [cycles/tone] given by an ellipse:

$$(f/W)^2 + (u_1/U_1)^2 = 1 \quad (6.54)$$

where  $W$  [Hz] and  $U_1$  [cycles/tone] give the highest tonal and temporal frequencies in  $\bar{\mathbf{d}}(k)$ . In our study, we used a sampling interval of  $T = 0.4s$  and the tonal range is  $A_1 = 1$ . With  $M_t = M_1 = 16$ , this gives a tonal temporal Nyquist frequencies of  $(f_N, u_N) = (0.078 \text{ [Hz]}, 8.0 \text{ [cycles/tone]})$ .

We consider two cases of sampling: full sampling where all  $M_t$ -measurement points are used at each sampling instant,  $k$  and time-sequential sampling where only  $n = 1$  tone is sampled at each sampling instant according to a prescribed sampling pattern i.e. lexicographical or bit-reversed sampling sequences. As our primary interest is to analyze the effect of disturbances with different sampling schemes, we assume the measurement noise  $\mathbf{n}(k) \approx 0$ .

The LQ with integral dynamics controller that minimizes (6.4) with one ( $q_t = 1$ ) integral fix point at tone,  $x_2$  is obtained from (6.6). The performance weighting matrices  $\mathbf{W}_s$  and  $\mathbf{W}_e$  in (6.4) are identities. The time-normalized quadratic performance index,  $\text{QPI}_L = \sum_{k \in [l_0, l_0+L]} J(k)/L$  is taken as the measure of performance of the TRC stabilization controller.

Simulations were carried out at different tonal-temporal support frequencies  $(W, U_1)$  for  $\bar{\mathbf{d}}(k)$  within the range of  $\{(W, U_1) \mid 0.01 \leq W \leq 2f_N, 1 \leq U_1 \leq u_N\}$  [cycles/tone,Hz] using the LQ TRC stabilization controller. As shown in Figure 6.11, both the (lexicographical) time-sequential sampling and the full sampling have comparable  $\text{QPI}_L$  TRC error (less than 0.10) for the range of input frequencies tested. This means that we are able to achieve as good TRC stabilization by just sampling one TRC tone at each print cycle, as by sampling all  $M_t = 16$  tones. Figure 6.11 also shows the expected response with increasing tonal-temporal support frequencies of the disturbance signal – the higher the tonal-temporal support frequencies, the higher the  $\text{QPI}_L$  TRC error. The higher tonal-temporal disturbance

frequencies, the harder it is for the TRC controller to curve-fit the measured TRC to the desired TRC.

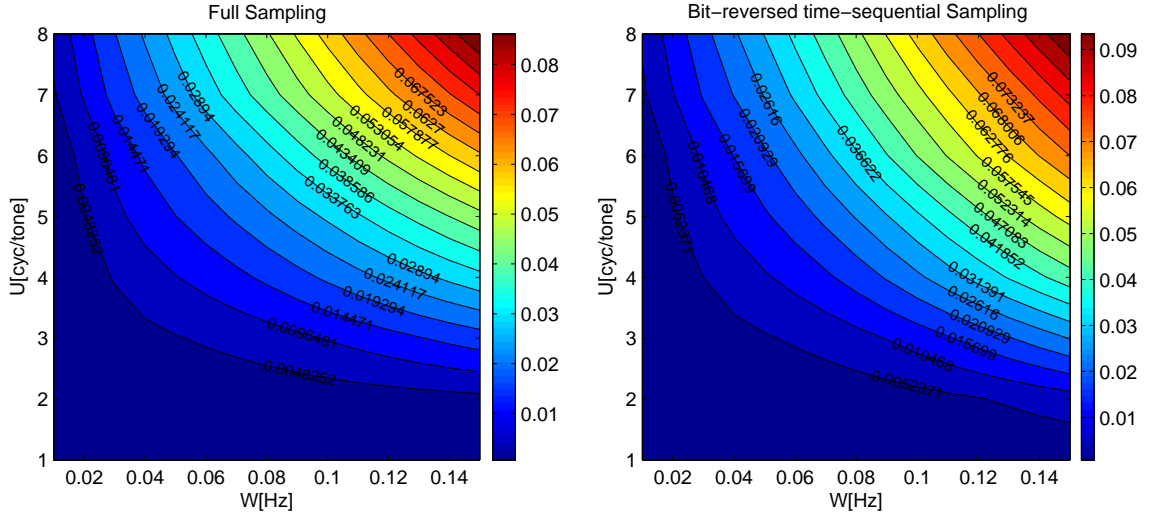


Figure 6.11:  $QPI_L$  of TRC error using full and lexicographic time-sequential sampling at different tonal-temporal frequencies support  $(W, U_1)$

For the simulated range of  $(U_1, W)$  values, control using the bit-reversed time-sequential sampling sequence out-performs control using the lexicographic sequence as indicated by the lower  $QPI_L$  TRC error for the bit-reversed sampling sequence than that of the lexicographic sequence (maximum of 19.5% improvement). This is a consequence of the better signal reconstruction of the bit-reversed sampling sequence as reported in Chapter 3.

Figure 6.12 shows the effectiveness of the proposed TRC stabilization system using both full and time-sequential samplings in an experimental setting. The disturbances were artificially induced by introducing a transparency in the optical path of the laser. Sensing of the color wedge was performed using a scanner. The TRC stabilization using all the sampling approaches results in the convergence of the TRC to the desired TRC with each print cycle. Considering that only one tone is sampled at each print cycle, the time-sequential sampling perform well compared to that achieved using full sampling ( $M_t = 16$ ).

## Case Study 2

From **Case Study 1** we have shown that:

1. The TRC stabilization control system can be implemented practically and the control performance with one time-sequential sampled sensor is nearly as effective as using full sampling with sixteen fixed sensors.

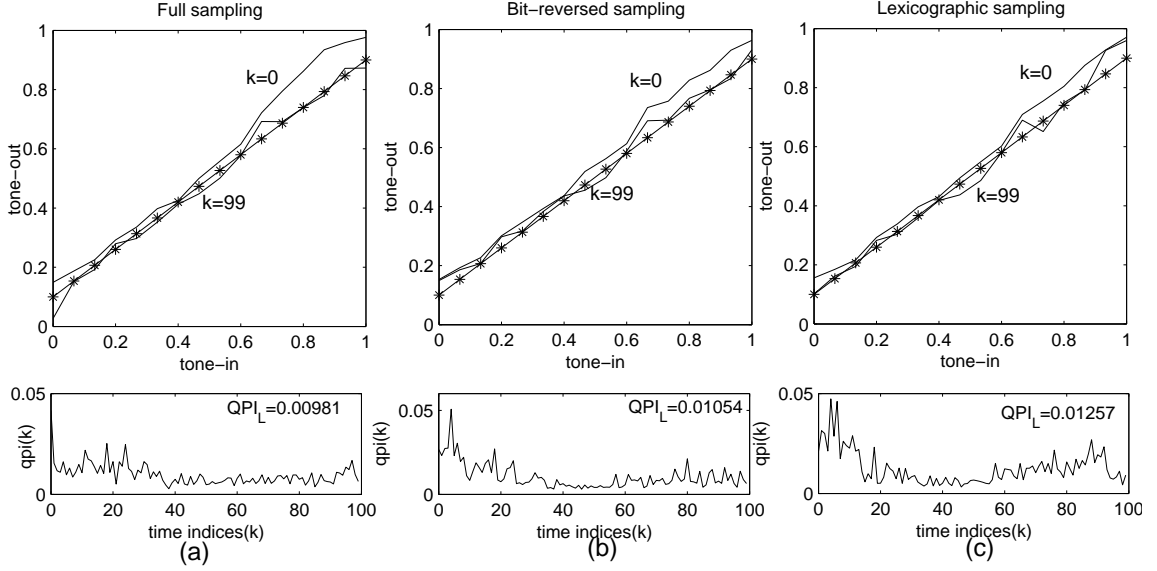


Figure 6.12: Response of TRC stabilization control subjected to induced disturbance with; (a) full sampling; (b) bit-reversed  $TS(1)$ ; (c) Lexicographic  $TS(1)$ . The curve marked with \* is the desired TRC . The  $QPI_L$  at steady-state, computed with  $l_0 = 50$ ,  $L = 50$  for the full, bit-reversed and lexicographic sampling is given by 0.00981, 0.01054, 0.01257 respectively.

2. Effectiveness of the TRC stabilization control depends on the accuracy of the signal reconstruction.

In this case study, experimental study was conducted for both STAGE-I and STAGE-II control for the case of the TRC using the optimal  $TS(1)$  sampling approach for all the CMY color separations. The objective of this study is to validate the proposed TRC stabilization control approach in achieving the control objective as given in (6.1). The following specific system behaviors were considered:

- (i) The behavior of the system without plant perturbation is considered i.e.  $\Delta(k) = \mathbf{0}$ .
- (ii) The disturbance  $\bar{\mathbf{d}}(k)$  dynamics is simulated from the model given in (6.53). The assumed spectral support  $\Theta(W, U_1)$  is defined by  $W = 0.04$  [Hz],  $U_1 = 3.0$  [cycles/tone] as shown in Figure 3.1. The low-frequency band dynamics is given by (3.12) with the temporal cutoff frequency  $cf_1 = 0.05W$  [Hz] and white process noise covariance,  $\mathbf{R}_1 = 0.01\mathbf{I}$ . The high-frequency band dynamics is given by (3.13) with the temporal cutoff frequency,  $f$  of the 10-th order low-pass filter at each spatial channel,  $u_1$  given by an ellipse:

$$(u_1/\bar{U}_1)^2 + (f/\bar{W})^2 = 1$$

where  $\bar{U}_1$  [cycles/tone] and  $\bar{W}$  [Hz] give the highest tonal and temporal frequencies in  $\bar{\mathbf{d}}(k)$ . In our study, we assume that  $\bar{U}_1 = 0.8U_1$  [cycles/tone] and  $\bar{W} = 0.8W$  [Hz] to account for the frequency drop-off characteristics of the non-ideal low-pass filter. Also, the white process noise covariance for the high-frequency band dynamics given by  $\mathbf{R}_2 = 0.02\mathbf{I}$ .

By the optimal  $TS(1)$  design given in section 5.6, (with  $A_1 = 1, M_{max,1} = 50$ ) gives an inter-sampling time  $T = T^* = 2.28\text{s}$  and  $M_1 = M^* = 39$ .

The LQ with integral dynamics controller that minimizes (6.4) with two ( $q_t = 2$ ) integral fix point at input tones  $x_1$  and  $x_{M_t}$  is obtained from (6.6) to maintain the range of the TRC for STAGE-II compensation. The performance weighting matrices  $\mathbf{W}_s$  and  $\mathbf{W}_e$  in (6.4) are identities. The time-normalized mean square error:

$$\text{MSE}_L = \left[ \frac{1}{L} \sum_{k \in [l_0, l_0+L]} \|\bar{\mathbf{e}}(k)\|_2^2 \right]^{1/2}$$

is taken as the measure of performance of the TRC stabilization controller.

Figure 6.13-6.15 show the effectiveness of the TRC stabilization control using the TRC estimates by full sampling, 3 fixed points sampling and optimal  $TS(1)$  sampling for the CMY primaries respectively. The proposed TRC stabilization control using the optimal  $TS(1)$  sampling is shown to be effective in compensating for the injected disturbances.

Figure 6.16 shows the  $\Delta E_{94}^*$  of the CMY image map showing the quantitative stabilization performance of the TRC stabilization control using the estimates by optimal  $TS(1)$  sampling and 3 fixed points sampling. The performance using the estimate by optimal  $TS(1)$  sampling is better than that using the 3 fixed points sampling (also shown in Figure 6.13), while requiring 3 times less samples. In our experiment, a calibration step is required to characterize the inverse print map which induced undue sensing noise into the sensing system. Hence, sensing noise in this experiment is higher than if we can directly sense the TRC (see discussion in section 5.6.4). The effectiveness in maintaining the color consistency of the proposed color stabilization strategy is shown by a series of time lapse images shown in Figure 6.17.

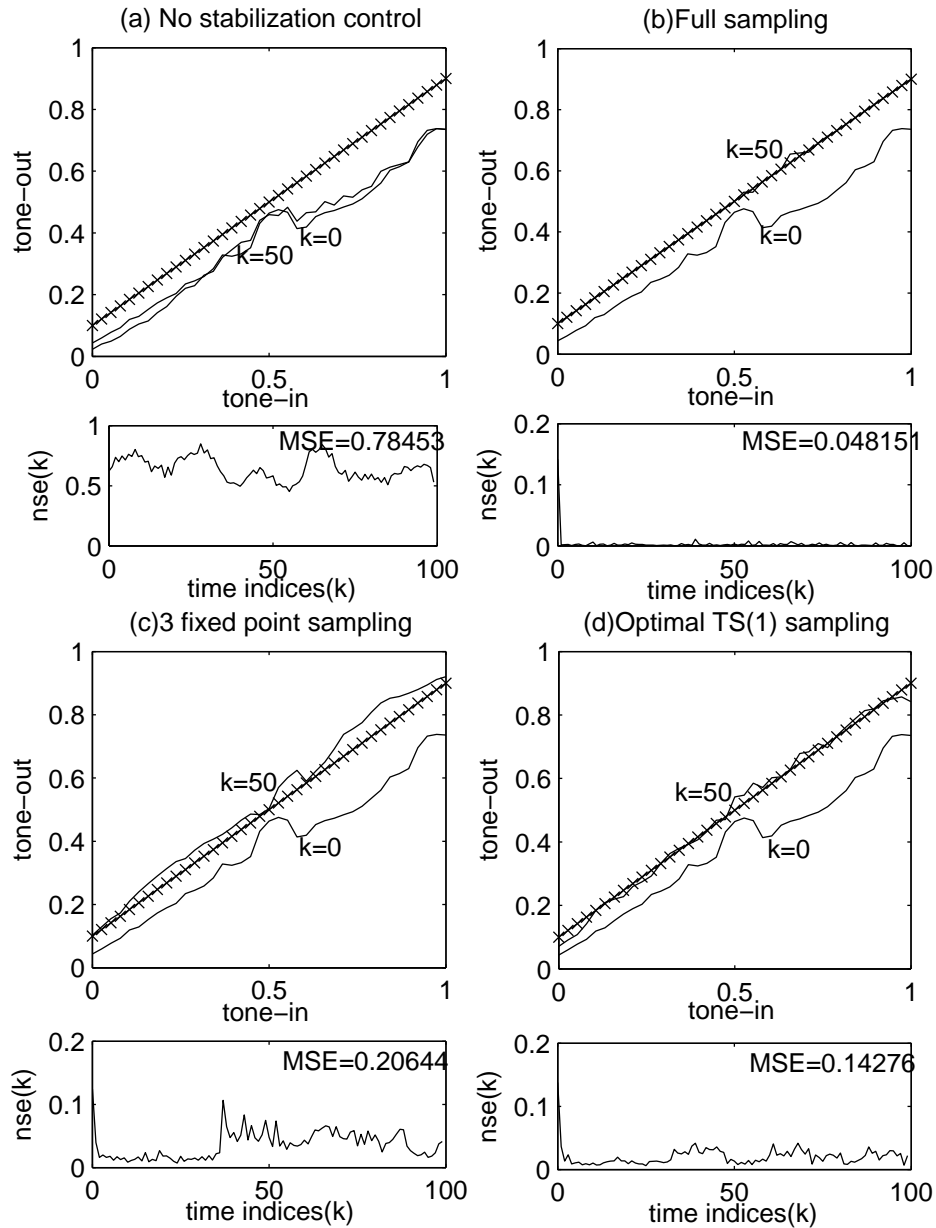


Figure 6.13: Response of TRC for Cyan primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal  $TS(1)$  sampling. The curve marked with 'x' is the desired TRC. The  $MSE_L$  at steady-state, computed with  $l_0 = 50, L = 50$  are shown at each respective figures.

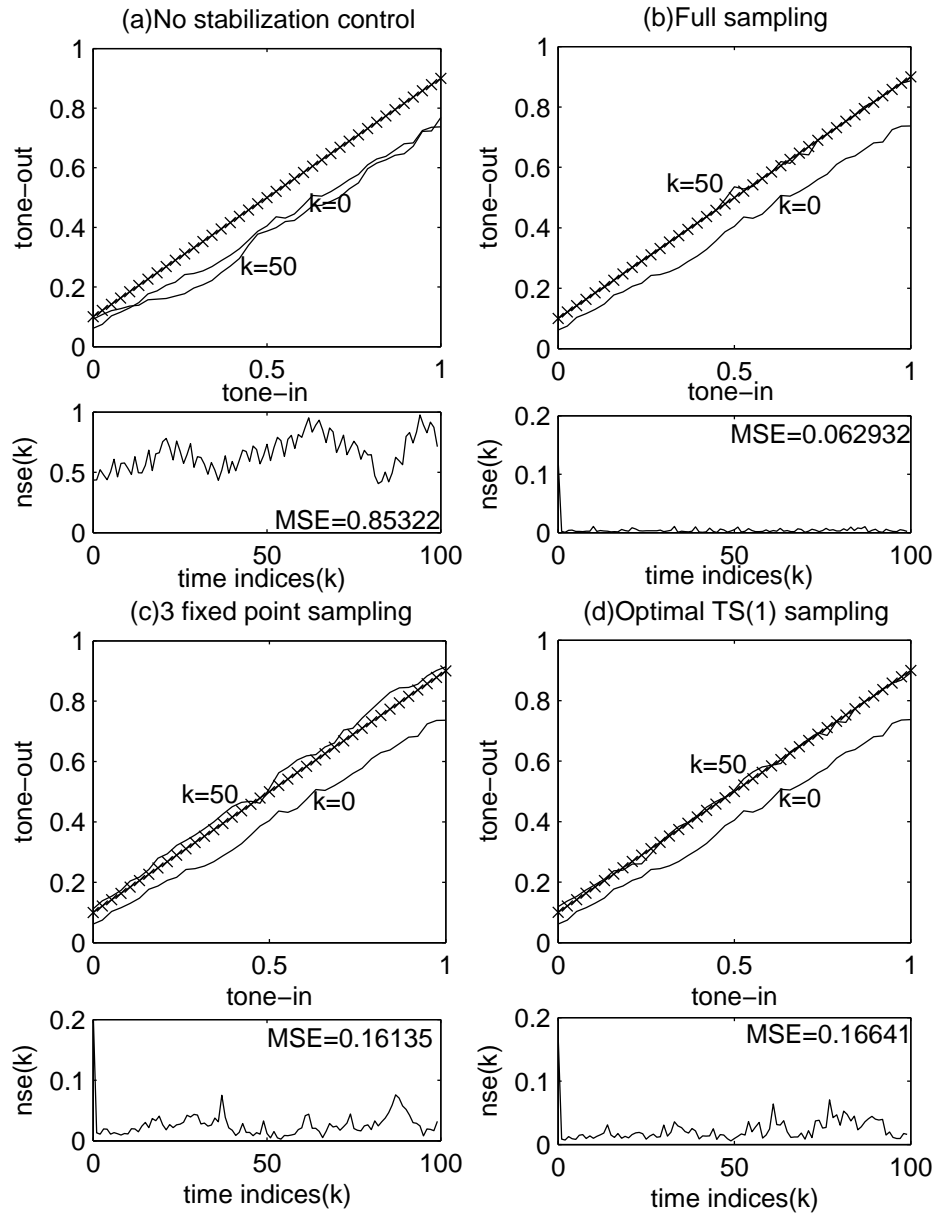


Figure 6.14: Response of TRC for Magenta primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal  $TS(1)$  sampling. The curve marked with 'x' is the desired TRC. The  $MSE_L$  at steady-state, computed with  $l_0 = 50, L = 50$  are shown at each respective figures.



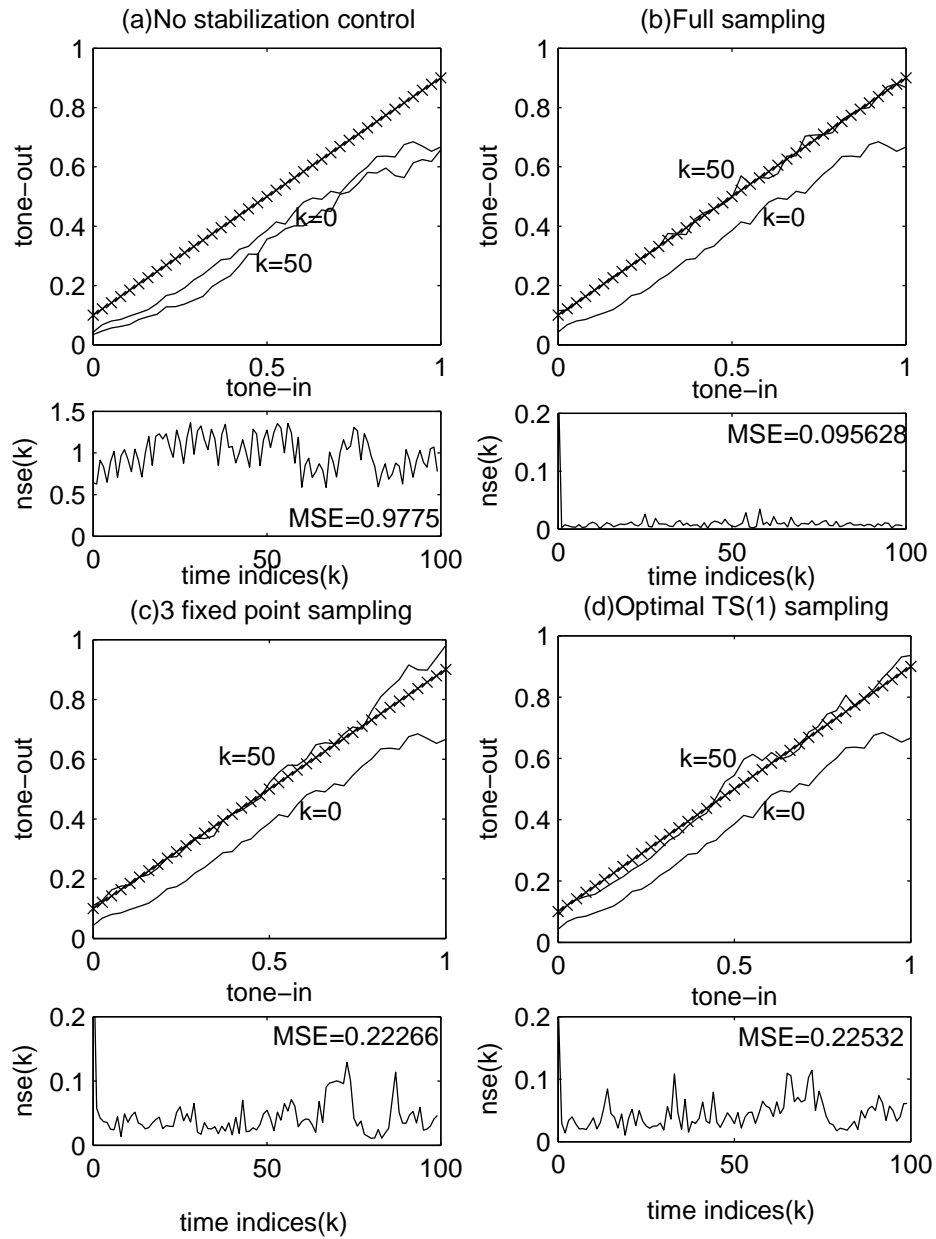


Figure 6.15: Response of TRC for Yellow primary subjected to induced disturbance with; (a) no stabilization control; (b) stabilization control with full sampling; (c) stabilization control with 3 fixed point sampling; (d) stabilization control with optimal  $TS(1)$  sampling. The curve marked with 'x' is the desired TRC. The  $MSE_L$  at steady-state, computed with  $l_0 = 50, L = 50$  are shown at each respective figures.

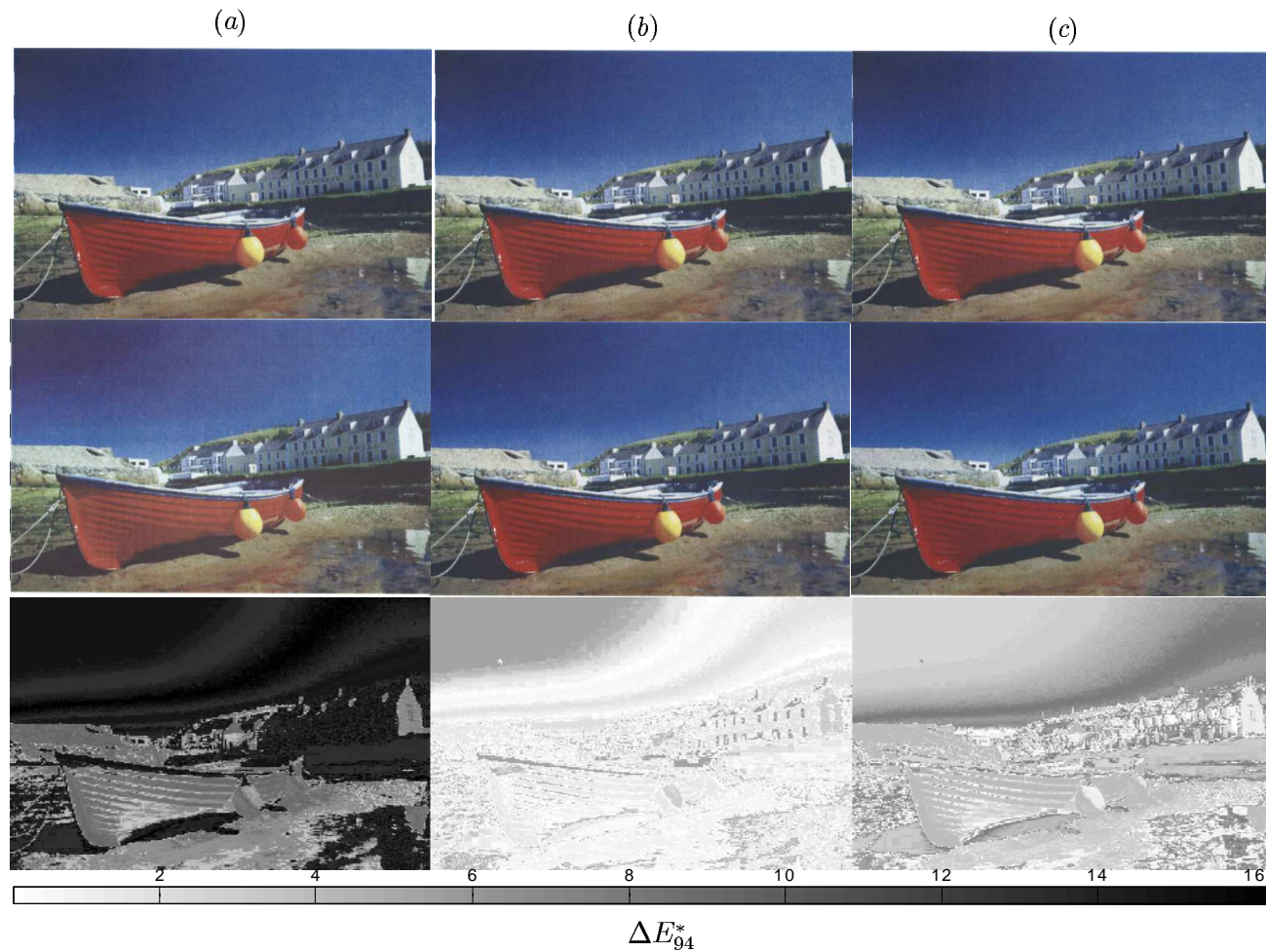


Figure 6.16: The reference image (top row images), the printed time-varying image (middle row images) with; (a) no TRC stabilization control; (b) TRC stabilization control using TRC estimate by optimal  $TS(1)$  sampling; (c) TRC stabilization control using TRC estimate by 3 fixed point sampling, at print cycle  $k = 50$ . The  $\Delta E_{94}^*$  map of the  $CMY$  image comparing the top row images and middle row images are given at the bottom rows figures

### 6.3.2 CRC stabilization control

The proposed CRC stabilization system using both **STAGE-I** and **STAGE-II** control was experimentally tested on a Xerox Phaser 7700 xerographic printer. A X-rite DTP70 scanning spectrophotometer was used to measure the printed color patches. The objective of this study is to validate the effectiveness of the proposed sensing and control approaches to stabilize the CRC map. Figure 6.18 shows the schematic of the experimental setup. The following system behaviors were considered:

- (i) To evaluate the proposed CRC stabilization control system, virtual printer models were first used to generate the tonal responses of the CMY primaries. This is because we do not have direct access to the xerographic actuators. These virtual printer models are given by the form described by (6.3) for the case of the TRC taking no disturbance, and no plant perturbation. The nominal sensitivity functions for each primary color separation is similar to that used for the TRC stabilization control. Motivated by the Yule-Nielsen Spectral Neugebauer model (shown in Figure 2.5), the CMY gamut described by the TRC for each of the primary CMY colorant is input to the Xerox Phaser 7700 color printer as shown in Figure 6.18. This form our physical color printing setup.
- (ii) The disturbances were artificially introduced by injecting a simulated disturbance source as shown in Figure 6.18 given by  $f_{P,0}^{-1} \mathbf{d}_{inject}(k)$  where  $f_{P,0}^{-1} : \mathcal{C}_{L^*a^*b^*} \rightarrow \mathcal{C}_{CMY}$  at nominal and  $\mathbf{d}_{inject}(k)$  is simulated from the model given in (6.53), with the assumed spectral support  $\Theta(W, \mathbf{U})$  where  $W = 0.01$  [Hz] and  $\mathbf{U} = [2, 2, 2]$  [cycles/toner]. With this disturbance model, the low-frequency band dynamics is given by (3.12) with the temporal cutoff frequency  $cf_1 = 0.05W$  and covariance matrix  $\mathbf{R}_1 = 0.001\mathbf{I}$ . The high-frequency band dynamics is given by (3.13) with the temporal cutoff frequency,  $f$  of the 4-th order low-pass filter at each spatial channel,  $\mathbf{u}$  given by an ellipse,

$$\Theta(\bar{W}, \bar{\mathbf{U}}) = \left\{ (f, \mathbf{u}) \left| f^2/\bar{W}^2 + \sum_{i=1}^N u_i^2/\bar{U}_i^2 = 1 \right. \right\}$$

where  $\bar{\mathbf{U}} := [\bar{U}_1, \dots, \bar{U}_N] = 0.8\mathbf{U}$  [cycles/toner] and  $\bar{W} = 0.8W$  [Hz]. A scaled factor of the actual support size is used here to account for the frequency drop off characteristic of the non-ideal low-pass filter used to approximate the disturbances. The white process noise covariance is given  $\mathbf{R}_2 = 100\mathbf{I}$ . This completes our hardware-in-loop xerographic color printing system.

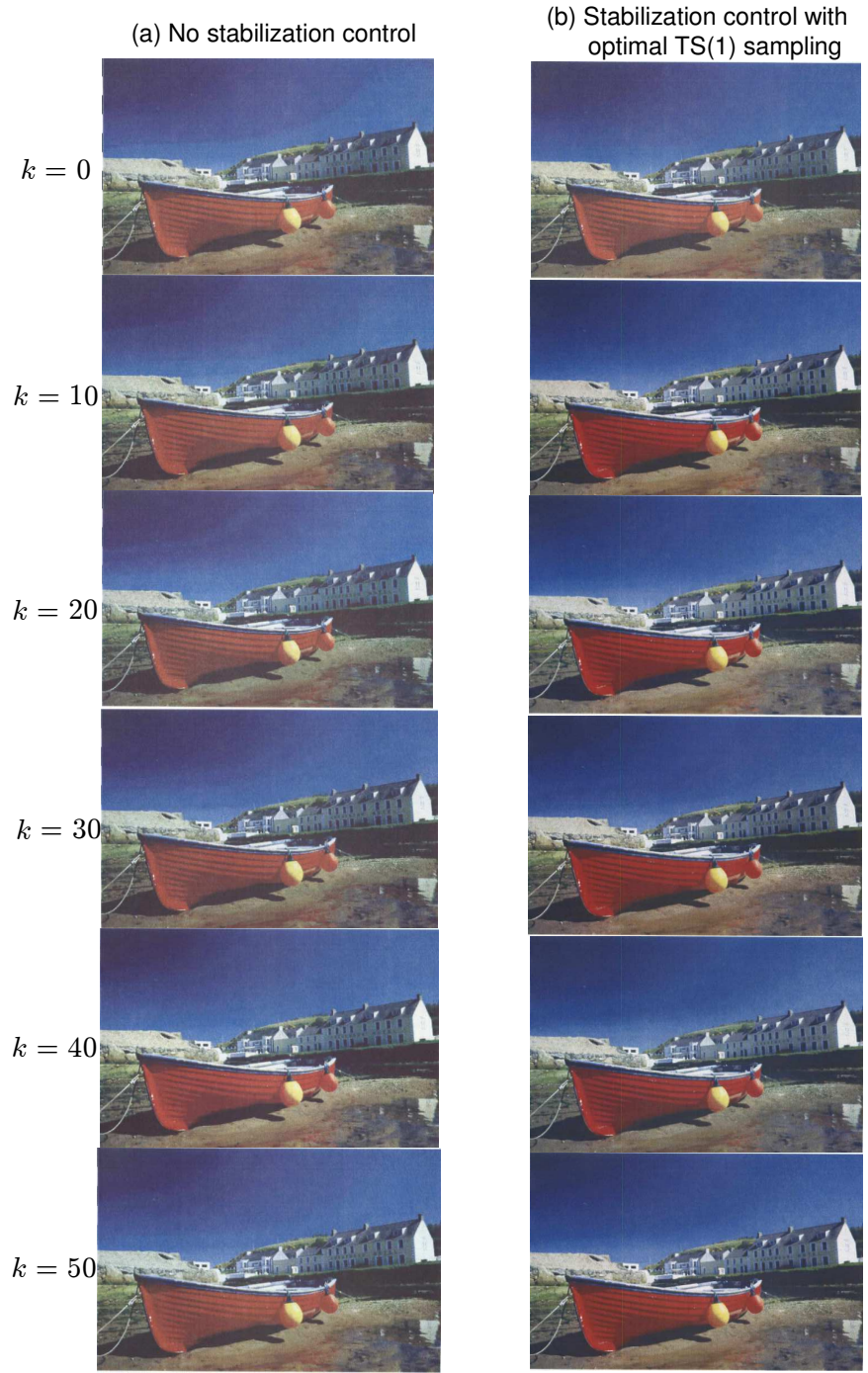


Figure 6.17: Color consistency with (right) and without (left) color consistency control at different print cycles

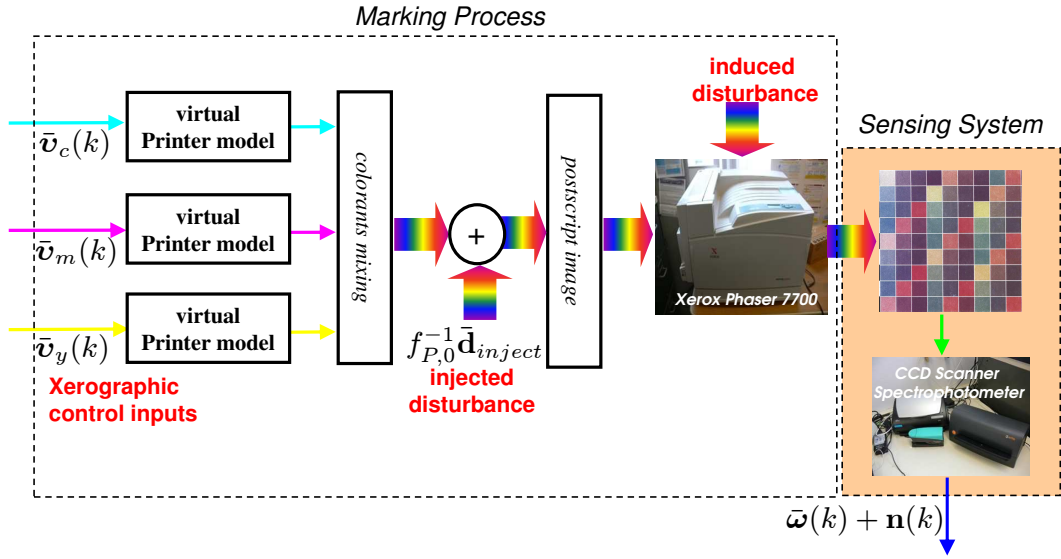


Figure 6.18: Experimental xerographic marking setup for CRC stabilization

While with this setup, the compact spectral support of the time-varying CRC is well-known, it is necessary to determine this in the actual system. For the actual system, determination of the compact spectral support is given by the procedure described in Appendix B.2. Hence, it is now possible to design the optimal  $TS(1)$  sampling as given in section 5.6 for this particular case of spectral support. Knowledge of the spectral support and the underlying time-varying CRC data (for deriving the color basis matrix) also allows for the design of the appropriate periodic Kalman filter as given in section 3.3. This means that we have the desired feedback information, obtained from the reconstructed time-varying CRC,  $\hat{\boldsymbol{\omega}}(k)$  at each print cycle,  $k$ .

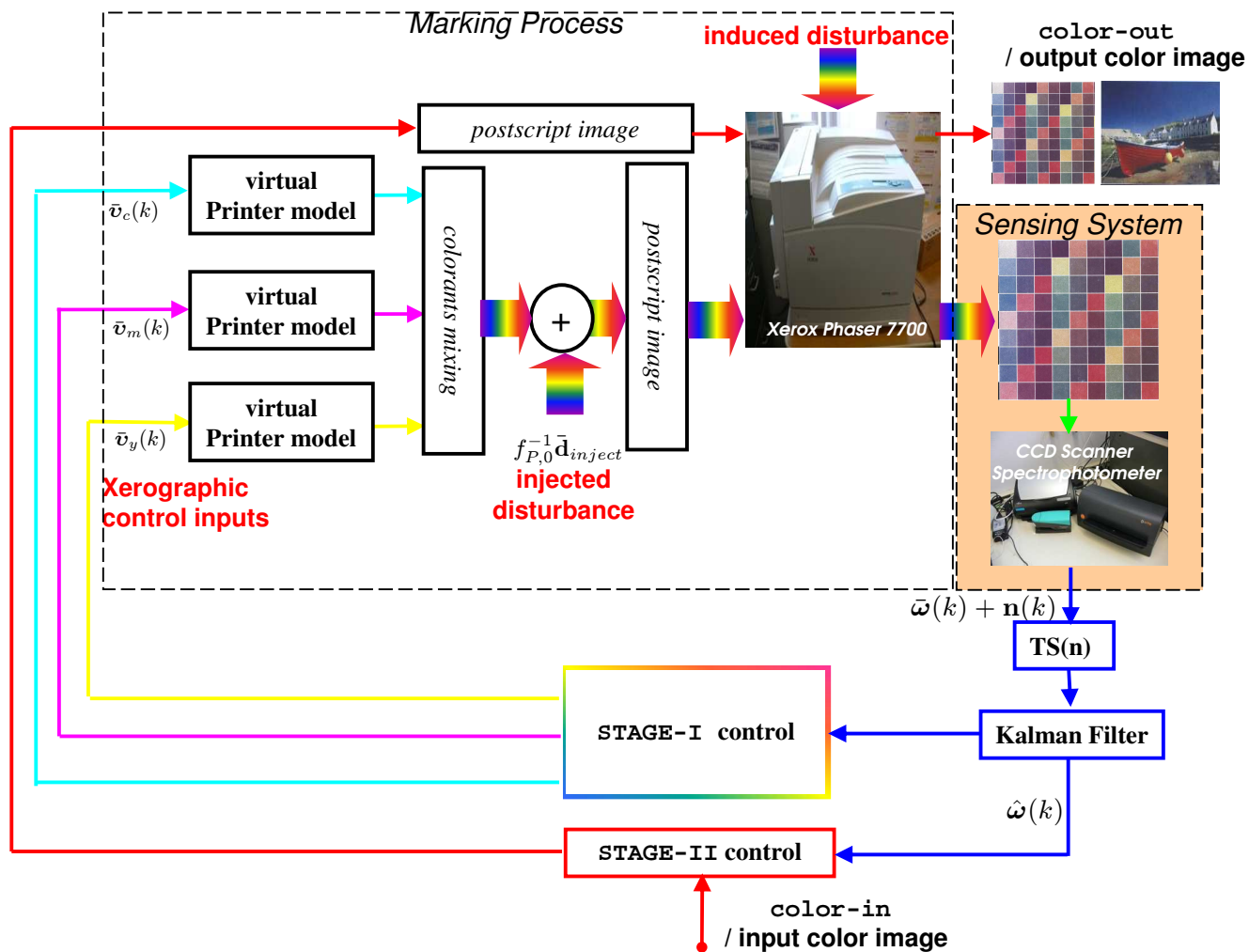


Figure 6.19: Sensing and control for CRC stabilization

To design the **STAGE-I** controllers as proposed in this chapter, requires the identification of the nominal sensitivity function  $\hat{\phi}$  for the case of the CRC. This is achieved by performing factorial experiments. In this data each experiment corresponds to each of the actuator settings of the CMY primary color separation at one of the 2 settings within their range. With  $m = 3$  actuators, we have a total of  $2^3 = 8$  different settings for each of the primary. With the CMY primary color separations we have a total of  $8^3 = 512$  different set of actuator settings to consider. From these data, the nominal sensitivity function,  $\hat{\phi}$  were then obtained using the least square method. **STAGE-I** maintains the range of the CRC with residual variations compensated in the image processing process using **STAGE-II** control. In **STAGE-II** control, the inverse print map is achieved with interpolation using the multidimensional look-up table approach [34]. In this experiment the nominal input gamut,  $\bar{\mathcal{C}}_{L^*a^*b^*}$  for **STAGE-II** compensation is selected such that  $\bar{\mathcal{C}}_{L^*a^*b^*} \subset \hat{\mathcal{C}}_{L^*a^*b^*}$  at all times to avoid infeasible interpolation at the boundary. The color consistency control system setup is shown in Figure 6.19.

Figure 6.20 shows the CRC stabilization performance of the linear quadratic with integral control as the **STAGE-I** control with the optimal  $TS(1)$  sampling approach. The performance weighting matrices in (6.4) are taken as identity for  $\mathbf{W}_s$  with one arbitrarily selected fixed point and  $\mathbf{W}_e$  is obtained from procedures given in section 6.1.3. Figure 6.21 shows the CRC stabilization performance of the robust static control as the **STAGE-I** control with the optimal  $TS(1)$  sampling approach. The performance weighting matrices in (6.22) are taken as identity for  $\bar{\mathbf{W}}_v$  and  $\bar{\mathbf{W}}_e$  is obtained from procedures given in section 6.1.3. Compared to the case without any stabilization control, the **STAGE-I** control effectively reduces the error between the printed output color and the nominal color gamut. Note however that the effectiveness of **STAGE-I** control depends largely on the degree in which the disturbance signal is in the range space of the sensitivity function,  $\hat{\phi}$ .

More effective stabilization can be achieved using **STAGE-II** control as shown in both Figure 6.20 and Figure 6.21. In addition, Figure 6.22 shows that the resulting CRC stabilization performance is better using the optimal  $TS(1)$  sampling compared to the lexicographic  $TS(1)$  sampling in sensing the time-varying CRC. This result indicates the importance of requiring good estimates of the time-varying CRC for successful implementation of the two-stage control strategy - a requirement that is effectively and efficiently fulfilled using the proposed  $TS(n)$  sampling approach.

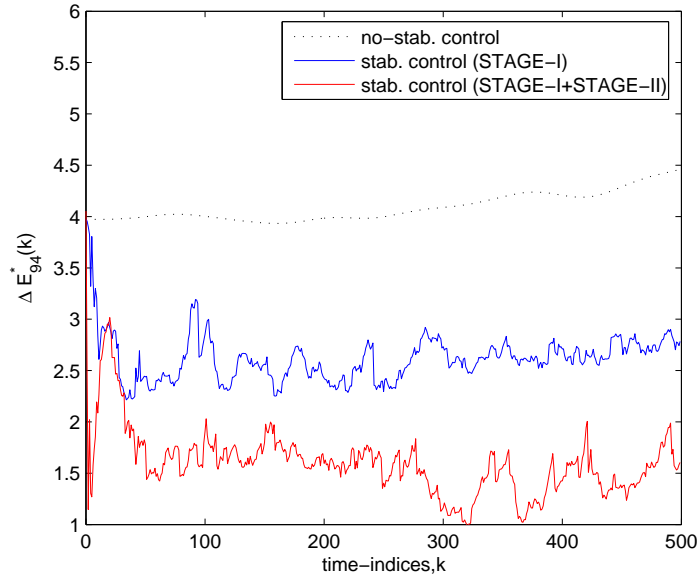


Figure 6.20: Comparing CRC stabilization performance using LQ with integral control for STAGE-I and STAGE-II compensation with the case of no stabilization control.

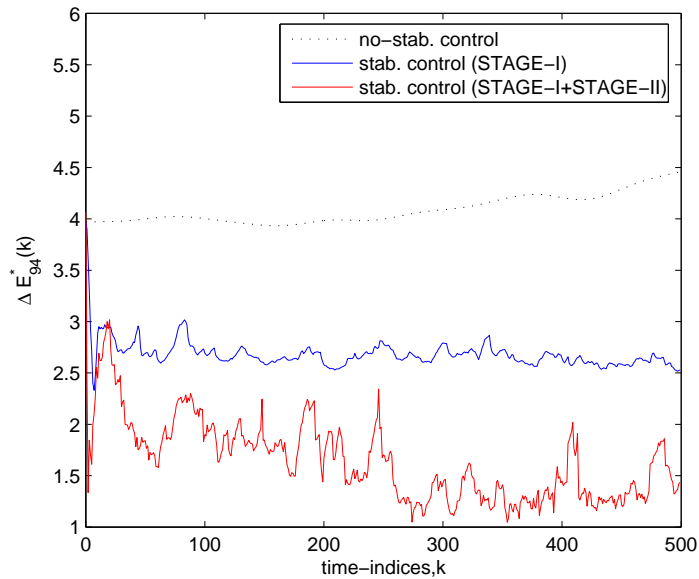


Figure 6.21: Comparing CRC stabilization performance using robust static control for STAGE-I and STAGE-II compensation with the case of no stabilization control



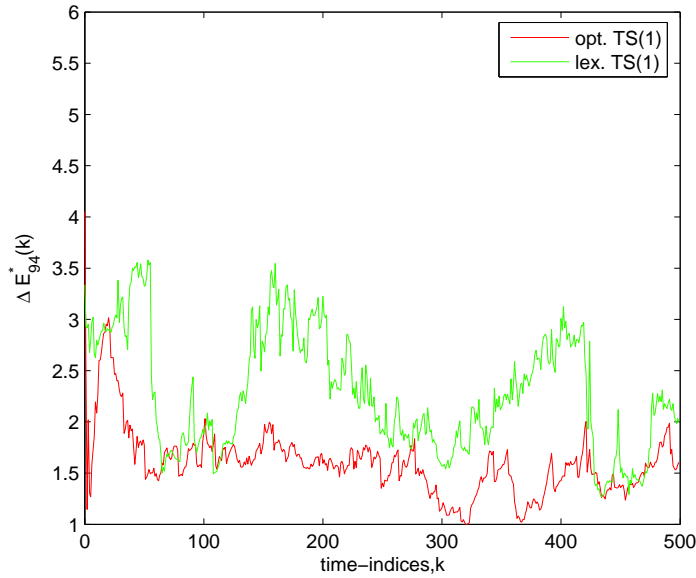


Figure 6.22: Comparing CRC stabilization performance using optimal  $TS(1)$  sampling and lexicographic  $TS(1)$  sampling. The linear quadratic control is used as **STAGE-I**

#### 6.4 Summary

In this dissertation tone reproduction curve (TRC) and color reproduction characteristics (CRC) function stabilization controllers for maintaining tone and color consistency are proposed. The control objective is to control the mapping induced by the xerographic process i.e. **STAGE-I** control via full closed loop feedback, and the image processing processes via feed-forward control i.e. **STAGE-II** control such that on nominal values of the human visual perception model, the TRC/CRC is close to identity. The **STAGE-I** control makes use of the xerographic actuators to ensure that the range of the TRC/CRC is preserved. Then, a **STAGE-II** compensation adjusts the image processing step such that the TRC/CRC is as close as possible to the desired TRC/CRC. Both simulation and experimental results show the effectiveness of the proposed approach for TRC/CRC stabilization control. In particular the use of optimal  $TS(1)$  sampling enables better TRC/CRC stabilization performance compared to the use of the conventional 3 fixed point sampling or the lexicographic  $TS(1)$  sampling.

## Chapter 7

# Color Coordination Control

In this chapter, the problem of coordinating multiple xerographic printers to *consistently* print the *same colors* is addressed. Coordination enables multiple printers to print the same image or collection of images in parallel, hence speeding the print process without resulting in inconsistent colors. This is an important development for recent technological development in digital color printing called *reconfigurable parallel printing* [20]. In this task maintaining printer-to-printer colors/tones consistency is as important as, or sometime more important than page-to-page consistency. For purpose of illustration and without loss of generality, consider two printers denoted as printer 1 and printer 2. From (2.4), the CRC map of both printers are given by:

$$CRC_1(t) = \underbrace{f_{percept}}_{\text{human}} \circ \underbrace{{}^1f_{comb} \circ {}^1f_{mark}}_{\text{printer}} \circ \underbrace{{}^1f_{htone} \circ {}^1f_{sep}}_{\text{image processing}} \quad (7.1)$$

$$CRC_2(t) = \underbrace{f_{percept}}_{\text{human}} \circ \underbrace{{}^2f_{comb} \circ {}^2f_{mark}}_{\text{printer}} \circ \underbrace{{}^2f_{htone} \circ {}^2f_{sep}}_{\text{image processing}} \quad (7.2)$$

The coordinated stabilization control strategy is to coordinate  ${}^1f_{comb}(t) \circ {}^1f_{mark}$  and  ${}^2f_{comb} \circ {}^2f_{mark}$  of printer 1 and printer 2 respectively, such that based on nominal values of  ${}^1f_{sep}$ ,  ${}^1f_{htone}$  and  ${}^2f_{sep}$ ,  ${}^2f_{htone}$  of printer 1 and printer 2 respectively, and the human visual perception model,  $f_{percept}$ , the mapping of both printers are close to identity.

If printer 1 and printer 2 are left uncoupled, a disturbance input applied to one of the two printers will cause color/tone variations in the disturbed printer only. Hence, prints from one printer to another may appear different although the page-to-page consistency is maintained. In this context we called these variations in color/tone as the coordination errors. To

achieve better coordination, a coupling controller which responds to the coordination error is introduced. In this case, when a disturbed input is applied to one of the printers, the color/tone variations appear in the undisturbed printer as well as in the disturbed printer. In doing so, we achieved better print consistency across different printers.

In this dissertation, only the coordination of the TRC at each  $j$ -th separation will be considered. The coordination control objective is for the TRC of printer 1 to match the TRC of printer 2 at each selected  $p_j$  tones,  $x_i \in \mathcal{T} \subset \mathbb{R}, i = 1, 2, \dots, p_j$ , as  $k \rightarrow \infty$  at each  $j$ -th separation

$$TRC_{1,j}(k)[x_i] \rightarrow TRC_{2,j}(k)[x_i] \quad (7.3)$$

**Remark 7.1**

1. *Similar to case of the TRC stabilization, if the manner in which the primary colors are combined is stable and constant, then the output colors will also be coordinated when the TRC for each separation has been effectively coordinated.*
2. *While it is true that if exact stabilization is achieved i.e. fulfilled objective given by (6.1) for each printer 1 and printer 2, then the coordination objective is similarly achieved, this is typically not attainable. This is because the number of tones to stabilize and coordinate is potentially high, and there are only a small number of effective actuators. Hence there may be loss of stabilization performance (i.e. page-to-page print consistency) with increase in coordination and vice versa. Therefore, emphasis must be given to stabilization or coordination as deemed necessary. In our case, coordination performance may sometime be more important than in stabilization because maintaining printer-to-printer prints consistency may be more desirable.*

In section 7.1 the coordinated control strategy is presented. An optimal static and robust static control approaches are proposed here enable for both the stabilization and coordination. Section 7.2 demonstrates the performance of the proposed coordination and stabilization control. Finally, section 7.3 gives some concluding remarks.

## 7.1 Coordinated Stabilization Control Strategy

In this section, without loss of generality, the coordinated stabilization control strategy is considered for a single primary separation. The high dimensionality of the TRC coupled with limited actuation authorities do not permit us to control all tones of both printers. Control approaches are proposed here to ensure that  $TRC_1(k)$  and  $TRC_2(k)$  are close to their nominal i.e.  $TRC_1^*$  and  $TRC_2^*$  respectively (for stabilization) and to each other (for

coordination) in a least squares sense. Augmenting the TRC errors from (6.3) of the two printers we have:

$$\begin{pmatrix} \bar{\mathbf{e}}_1(k) \\ \bar{\mathbf{e}}_2(k) \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{\phi}}_1 & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\phi}}_2 \end{pmatrix} \left( \mathbf{I} + \begin{pmatrix} \boldsymbol{\Delta}_1(k)\bar{\mathbf{W}}_{v1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta}_2(k)\bar{\mathbf{W}}_{v2} \end{pmatrix} \right) \begin{pmatrix} \bar{\mathbf{v}}_1(k) \\ \bar{\mathbf{v}}_2(k) \end{pmatrix} + \begin{pmatrix} \bar{\mathbf{d}}_1(k) \\ \bar{\mathbf{d}}_2(k) \end{pmatrix} \quad (7.4)$$

Let,

$$\bar{\mathbf{e}}(k) := \begin{pmatrix} \bar{\mathbf{e}}_1(k) \\ \bar{\mathbf{e}}_2(k) \end{pmatrix} \in \mathbb{R}^{(M_1+M_1)} \quad (7.5)$$

$$\hat{\boldsymbol{\phi}} := \begin{pmatrix} \hat{\boldsymbol{\phi}}_1 & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{\phi}}_2 \end{pmatrix} \in \mathbb{R}^{(M_1+M_1) \times (m_1+m_2)} \quad (7.6)$$

$$\begin{aligned} \boldsymbol{\Delta}(k)\bar{\mathbf{W}}_v &:= \begin{pmatrix} \boldsymbol{\Delta}_1(k) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta}_2(k) \end{pmatrix} \begin{pmatrix} \bar{\mathbf{W}}_{v1} & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{W}}_{v2} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{\Delta}_1(k)\bar{\mathbf{W}}_{v1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Delta}_2(k)\bar{\mathbf{W}}_{v2} \end{pmatrix} \in \mathbb{R}^{(m_1+m_1) \times (m_1+m_1)} \end{aligned} \quad (7.7)$$

$$\bar{\mathbf{v}}(k) := \begin{pmatrix} \bar{\mathbf{v}}_1(k) \\ \bar{\mathbf{v}}_2(k) \end{pmatrix} \in \mathbb{R}^{(m_1+m_1)} \quad (7.8)$$

$$\bar{\mathbf{d}}(k) := \begin{pmatrix} \bar{\mathbf{d}}_1(k) \\ \bar{\mathbf{d}}_2(k) \end{pmatrix} \in \mathbb{R}^{(M_1+M_1)} \quad (7.9)$$

In section 7.1.1 an optimal static control approach is developed by neglecting the system uncertainty  $\boldsymbol{\Delta}(k)$  in equation (7.4). Then, in section 7.1.2 an optimal robust control approach is developed that takes into account system uncertainties.

### 7.1.1 Optimal static stabilization and coordination controller

Consider the augmented system of (7.4) with  $\boldsymbol{\Delta}(k) = \mathbf{0}$ . Let  $\boldsymbol{\epsilon}(k)$  denotes the coordination errors given by:

$$\boldsymbol{\epsilon}(k) = \mathbf{H}_1\bar{\mathbf{e}}_1(k) - \mathbf{H}_2\bar{\mathbf{e}}_2(k) \quad (7.10)$$

where  $\mathbf{H}_1 \in \mathbb{R}^{p \times M_1}$  and  $\mathbf{H}_2 \in \mathbb{R}^{p \times M_2}$  give the interpolation matrices such that we pick collection of  $p$  tones to coordinate for printer 1 and printer 2 respectively. The optimal control problem is to find the control  $\bar{\mathbf{v}}(k) \in \mathbb{R}^{(m_1+m_2)}$  based on the measured TRCs such that the following quadratic performance index (QPI),  $J(k)$  is minimized.

$$J(k) = \frac{1}{2}\bar{\mathbf{e}}^T(k) \begin{pmatrix} \mathbf{W}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 \end{pmatrix} \bar{\mathbf{e}}(k) + \frac{1}{2}\bar{\mathbf{v}}^T(k) \begin{pmatrix} \mathbf{R}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_2 \end{pmatrix} \bar{\mathbf{v}}(k) + \frac{1}{2}\boldsymbol{\epsilon}^T(k)\mathbf{Y}\boldsymbol{\epsilon}(k) \quad (7.11)$$

where  $\mathbf{W}_1 \in \mathbb{R}^{M_1 \times M_1}$  and  $\mathbf{W}_2 \in \mathbb{R}^{M_2 \times M_2}$  are weighting matrices of the TRC errors for printer 1 and printer 2 respectively.  $\mathbf{R}_1 \in \mathbb{R}^{m_1 \times m_1}$  and  $\mathbf{R}_2 \in \mathbb{R}^{m_2 \times m_2}$  are weighting matrices of the xerographic actuators inputs.  $\mathbf{Y} \in \mathbb{R}^{p \times p}$  is the weighting matrix for the coordination of printer 1 and printer 2.

Substituting (7.4) and (7.10) into (7.11), we have:

$$J(k) = \frac{1}{2} \bar{\mathbf{v}}^T(k) \mathbf{Q}_1 \bar{\mathbf{v}}(k) + \frac{1}{2} \bar{\mathbf{d}}^T(k) \mathbf{Q}_2 \bar{\mathbf{d}}(k) + \frac{1}{2} \bar{\mathbf{v}}^T(k) \mathbf{Q}_3 \bar{\mathbf{d}}(k) + \frac{1}{2} \bar{\mathbf{d}}^T(k) \mathbf{Q}_3^T \bar{\mathbf{v}}(k) \quad (7.12)$$

where

$$\mathbf{Q}_1 = \begin{pmatrix} \hat{\phi}_1^T \mathbf{W}_1 \hat{\phi}_1 + \hat{\phi}_1^T \mathbf{H}_1^T \mathbf{Y} \mathbf{H}_1 \hat{\phi}_1 + \mathbf{R}_1 & -\hat{\phi}_1^T \mathbf{H}_1^T \mathbf{Y} \mathbf{H}_2 \hat{\phi}_2 \\ -\hat{\phi}_2^T \mathbf{H}_2^T \mathbf{Y} \mathbf{H}_1 \hat{\phi}_1 & \hat{\phi}_2^T \mathbf{W}_2 \hat{\phi}_2 + \hat{\phi}_2^T \mathbf{H}_2^T \mathbf{Y} \mathbf{H}_2 \hat{\phi}_2 + \mathbf{R}_2 \end{pmatrix} \quad (7.13)$$

$$\mathbf{Q}_2 = \begin{pmatrix} \mathbf{W}_1 + \mathbf{H}_1^T \mathbf{Y} \mathbf{H}_1 & -\mathbf{H}_1^T \mathbf{Y} \mathbf{H}_2 \\ -\mathbf{H}_2^T \mathbf{Y} \mathbf{H}_1 & \mathbf{W}_2 + \mathbf{H}_2^T \mathbf{Y} \mathbf{H}_2 \end{pmatrix} \quad (7.14)$$

$$\mathbf{Q}_3 = \begin{pmatrix} \hat{\phi}_1^T \mathbf{W}_1 + \hat{\phi}_1^T \mathbf{H}_1^T \mathbf{Y} \mathbf{H}_1 & -\hat{\phi}_1^T \mathbf{H}_1^T \mathbf{Y} \mathbf{H}_2 \\ -\hat{\phi}_2^T \mathbf{H}_2^T \mathbf{Y} \mathbf{H}_1 & \hat{\phi}_2^T \mathbf{W}_2 + \hat{\phi}_2^T \mathbf{H}_2^T \mathbf{Y} \mathbf{H}_2 \end{pmatrix} \quad (7.15)$$

with dimensions  $\mathbf{Q}_1 \in \mathbb{R}^{(m_1+m_2) \times (m_1+m_2)}$ ,  $\mathbf{Q}_2 \in \mathbb{R}^{(M_1+M_2) \times (M_1+M_2)}$  and  $\mathbf{Q}_3 \in \mathbb{R}^{(M_1+M_2) \times (M_1+M_2)}$ .

**Controller 7.1:** *Optimal static stabilization and coordination controller*

The optimal static stabilization and coordination controller is given by:

$$\bar{\mathbf{v}}(k) = \mathbf{K}_{opt} \bar{\mathbf{e}}(k) \quad (7.16)$$

where  $\mathbf{K}_{opt} := (\mathbf{I} + \mathbf{K} \cdot \hat{\phi})^{-1} \mathbf{K} \in \mathbb{R}^{(m_1+m_2) \times (M_1+M_2)}$  with  $\mathbf{K} := -\mathbf{Q}_1^{-1} \mathbf{Q}_3 \in \mathbb{R}^{(m_1+m_2) \times (M_1+M_2)}$

However this controller cannot be realized since  $\bar{\mathbf{e}}(k)$  is not available until  $\bar{\mathbf{v}}(k)$  is given. Instead let the realization of the controller be of the form:

$$\bar{\mathbf{v}}(k+1) = \mathbf{A}_{opt} \bar{\mathbf{v}}(k) + \mathbf{B}_{opt} \bar{\mathbf{e}}(k) \quad (7.17)$$

$(\mathbf{A}_{opt}, \mathbf{B}_{opt})$  are deduced using (6.27) and (6.28) respectively with  $\mathbf{A}_{opt} \leftarrow \mathbf{A}_\Delta$  and  $\mathbf{B}_{opt} \leftarrow \mathbf{B}_\Delta$  by having  $\mathbf{K}_\Delta \rightarrow \mathbf{K}_{opt}$  from (7.20)

**Derivation:**

Taking the derivative of (7.12) with respect to  $\bar{\mathbf{v}}^T(k)$  we have:

$$\frac{\partial J(k)}{\partial \bar{\mathbf{v}}^T(k)} = \mathbf{Q}_1 \bar{\mathbf{v}}(k) + \mathbf{Q}_3 \bar{\mathbf{d}}(k) \quad (7.18)$$

Then taking  $\partial J(k)/\partial \bar{\mathbf{v}}^T(k) = 0$  we have:

$$\bar{\mathbf{v}}(k) = -\underbrace{\mathbf{Q}_1^{-1} \mathbf{Q}_3}_{\mathbf{K}} \bar{\mathbf{d}}(k) \quad (7.19)$$

Substituting (7.4) with  $\Delta(k) = \mathbf{0}$  into (7.19) we have:

$$\bar{\mathbf{v}}(k) = \underbrace{(\mathbf{I} + \mathbf{K} \cdot \hat{\phi})^{-1} \mathbf{K}}_{\mathbf{K}_{opt}} \bar{\mathbf{e}}(k) \quad (7.20)$$

### 7.1.2 Robust static stabilization and coordination controller

From the augmented plant model in (7.4), note that  $\Delta(k)$  and  $\bar{\mathbf{d}}(k)$  are unknown. Let  $\bar{\mathbf{U}}(z) = \mathbf{K}(z) \bar{\mathbf{E}}(z)$  where  $\mathbf{K}(z)$  is some linear feedback controller to be specified. Define the error weighting  $\bar{\mathbf{W}}_{e1} \in \mathbb{R}^{M_1 \times M_1}$ ,  $\bar{\mathbf{W}}_{e2} \in \mathbb{R}^{M_2 \times M_2}$  which specify the relative importance of the TRC error at different tones and the coordination error weighting  $\bar{\mathbf{W}}_{c1} \in \mathbb{R}^{M_1 \times M_2}$  together with  $\bar{\mathbf{W}}_{c2} \in \mathbb{R}^{M_2 \times M_1}$  to specify the relative importance of the TRC coordination at different tones. From (7.10),  $\bar{\mathbf{W}}_{c1}$  and  $\bar{\mathbf{W}}_{c2}$  are given as follows:

$$\bar{\mathbf{W}}_{c1} = \alpha \mathbf{H}_1^T \mathbf{H}_2 \quad (7.21)$$

$$\bar{\mathbf{W}}_{c2} = \alpha \mathbf{H}_2^T \mathbf{H}_1 \quad (7.22)$$

where  $\alpha \in \mathbb{R}$ . Increasing  $\alpha$  enables greater emphasis on coordination.

Collectively the coordination and stabilization errors metric is given by:

$$\begin{pmatrix} \bar{\mathbf{W}}_{e1} \bar{\mathbf{e}}_1(k) + \bar{\mathbf{W}}_{c1} \bar{\mathbf{e}}_2(k) \\ \bar{\mathbf{W}}_{c2} \bar{\mathbf{e}}_1(k) + \bar{\mathbf{W}}_{e2} \bar{\mathbf{e}}_2(k) \end{pmatrix} = \bar{\mathbf{W}}_{ec} \bar{\mathbf{e}}(k) \quad (7.23)$$

where

$$\bar{\mathbf{W}}_{ec} = \begin{pmatrix} \bar{\mathbf{W}}_{e1} & \bar{\mathbf{W}}_{c1} \\ \bar{\mathbf{W}}_{c2} & \bar{\mathbf{W}}_{e2} \end{pmatrix} \in \mathbb{R}^{(M_1+M_2) \times (M_1+M_2)} \quad (7.24)$$

The closed loop system can be expressed as a linear fractional transformation (LFT). Specif-

ically, we have:

$$\begin{pmatrix} \mathbf{w} \\ \bar{\mathbf{W}}_{ec}\bar{\mathbf{e}} \\ \bar{\mathbf{e}} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{0} & \mathbf{0} & \bar{\mathbf{W}}_v \\ \bar{\mathbf{W}}_{ec}\hat{\phi} & \bar{\mathbf{W}}_{ec} & \bar{\mathbf{W}}_{ec}\hat{\phi} \\ \hat{\phi} & \mathbf{I} & \hat{\phi} \end{pmatrix}}_{\mathbf{P}} \begin{pmatrix} \mathbf{v} \\ \bar{\mathbf{d}} \\ \bar{\mathbf{v}} \end{pmatrix} \quad (7.25)$$

with feedback connection  $\mathbf{v}(k) = \mathbf{\Delta}(k)\mathbf{w}(k)$  and  $\bar{\mathbf{U}}(z) = \mathbf{K}(z)\bar{\mathbf{E}}(z)$ .

Because of the static nature of the xerographic process, and the xerographic disturbances are generally slowly varying, the performance optimization is restricted to the steady-state case. Hence, since  $\bar{\mathbf{W}}_{ec}\bar{\mathbf{e}}^\infty$  is linear with respect to  $\bar{\mathbf{d}}^\infty$ , there exists some matrix,  $\mathbf{F}(\mathbf{P}, \mathbf{\Delta}^\infty, \mathbf{K}^\infty)$  such that

$$\bar{\mathbf{W}}_{ec}\bar{\mathbf{e}}^\infty = \mathbf{F}(\mathbf{P}, \mathbf{\Delta}^\infty, \mathbf{K}^\infty) \cdot \bar{\mathbf{d}}^\infty \quad (7.26)$$

The goal here is to find a controller  $\mathbf{K}(z)$  such that for the worst case performance and for as large a class of uncertainty  $\mathbf{\Delta}(\cdot)$  as possible,  $\|\mathbf{F}(\mathbf{P}, \mathbf{\Delta}^\infty, \mathbf{K}^\infty)\|_2$  is minimized. This is achieved by optimizing the following steady-state performance index:

$$\bar{\gamma}(\mathbf{K}^\infty) = \min \left\{ \gamma \left| \sup_{\|\mathbf{\Delta}^\infty\| \leq \frac{1}{\gamma}} \bar{\sigma}(\mathbf{F}(\mathbf{P}, \mathbf{\Delta}^\infty, \mathbf{K}^\infty)) \leq \gamma \right. \right\} \quad (7.27)$$

where  $\bar{\sigma}[\cdot]$  denotes the maximum singular value of its argument. Hence, the optimal controller dc gain is

$$\mathbf{K}_{rob} := \arg \min_{\mathbf{K}^\infty} \bar{\gamma}(\mathbf{K}^\infty) \quad (7.28)$$

Finding  $\mathbf{K}_{rob}$  can be achieved by procedure described in [3, 56].

**Controller 7.2:** *Robust static stabilization and coordination controller*

The robust static stabilization and coordination controller is given by:

$$\bar{\mathbf{v}}(k) = \mathbf{K}_{rob}\bar{\mathbf{e}}(k) \quad (7.29)$$

Similar to the case of the optimal static controller given by (7.20), this controller cannot be realized since  $\bar{\mathbf{e}}(k)$  is not available until  $\bar{\mathbf{v}}(k)$  is given. Instead let the realization of the controller be of the form:

$$\bar{\mathbf{v}}(k+1) = \mathbf{A}_{rob}\bar{\mathbf{v}}(k) + \mathbf{B}_{rob}\bar{\mathbf{e}}(k) \quad (7.30)$$

$(\mathbf{A}_{rob}, \mathbf{B}_{rob})$  are deduced using (6.27) and (6.28) respectively with  $\mathbf{A}_{rob} \leftarrow \mathbf{A}_\Delta$  and  $\mathbf{B}_{rob} \leftarrow \mathbf{B}_\Delta$ , by having  $\mathbf{K}_\Delta \rightarrow \mathbf{K}_{rob}$ .

## 7.2 Performance of the Coordination and Stabilization Control

The following behavior for the plant model given by (7.4) is considered:

- (i) We assume here that  $\widehat{\phi}_1 = \widehat{\phi}_2$ , and taken to be the same as the sensitivity function for the case of the TRC used in section 6.3 for each of the CMY primary separation. We consider  $M_1 = M_2 = 39$  number of discretized tones, and  $m_1 = m_2 = 3$  number of effective xerographic actuators. The number of tones to be coordinated is given by  $p = 10$
- (ii) The uncertainties of the following form are considered:

$$\Delta_1(k)\bar{\mathbf{W}}_{v1} = \delta_1\Delta_1 \quad (7.31)$$

$$\Delta_2(k)\bar{\mathbf{W}}_{v2} = \delta_2\Delta_2 \quad (7.32)$$

where  $\Delta_1(k)$  and  $\Delta_2(k)$  are randomly chosen with  $\|\Delta_1\| = 1$  and  $\|\Delta_2\| = 2$ .  $\delta_1 \in \mathbb{R}$  and  $\delta_2 \in \mathbb{R}$  are systematically varied. The weighting matrices are arbitrarily chosen to be  $\bar{\mathbf{W}}_{v1} = 0.1\mathbf{I}$  and  $\bar{\mathbf{W}}_{v2} = 0.1\mathbf{I}$ .

- (iii) The TRC disturbances,  $\bar{\mathbf{d}}_1(k)$  and  $\bar{\mathbf{d}}_2(k)$  of printer 1 and printer 2 respectively are modeled by the form given by (3.11). For printer 1, the low-frequency band dynamics is given by (3.12) with the temporal cutoff frequency  $cf_1 = 0.15W$  [Hz] and white process noise covariance of  $\mathbf{R}_1 = 0.03\mathbf{I}$ . The high-frequency band dynamics is given by (3.13) with the temporal cutoff frequency,  $f$  of a 10-th order low-pass filter at each spatial channel,  $u_1$  given by an ellipse,  $(u_1/\bar{U}_1)^2 + (f/\bar{W})^2 = 1$  where  $\bar{U}_1$  [cycles/toner] and  $\bar{W}$  [Hz] give the highest tonal and temporal frequencies in  $\bar{\mathbf{d}}_1(k)$ . In our study, we assume that  $\bar{U}_1 = 0.8U_1$  [cycles/toner],  $\bar{W} = 0.8W$  [Hz] to account for the frequency drop-off characteristics of the non-ideal low-pass filter. The white process noise covariance is selected to be  $\mathbf{R}_2 = 0.001\mathbf{I}$ . The parameters setting for printer 2 is the same except that a smaller cutoff frequency for the low-frequency band dynamics is selected i.e.  $cf_1 = 0.05W$  [Hz] in this case. This creates the required difference in disturbance in the two printers. A small sensing noise covariance is assumed for both printers i.e.  $\mathbf{R}_{nn} = 1.0 \times 10^{-4}\mathbf{I}$ .

All the  $M_1, M_2$  tones of the TRCs are available at each time-step,  $k$  either by using the  $TS(n_1)$  and  $TS(n_2)$  sampling approaches where small number of  $n_1$  and  $n_2$  tones are sampled and then reconstructed with a periodic Kalman filter as proposed in Chapter 3 to Chapter 5. The TRCs error norm,  $\|\bar{\mathbf{e}}_1(k)\|_2$  and  $\|\bar{\mathbf{e}}_2(k)\|_2$  and the coordination error norm,  $\|\boldsymbol{\epsilon}(k)\|_2$  are used as the measures of the the stabilization and coordination performance at



each time-step,  $k$  respectively. As measures of the overall stabilization and coordination performance, the time-normalized mean square TRCs error,

$$\text{MSE-STAB-1}_L = \left[ \frac{1}{L} \sum_{k \in [l_0, l_0+L]} \|\bar{\mathbf{e}}_1(k)\|_2^2 \right]^{\frac{1}{2}} \quad (7.33)$$

$$\text{MSE-STAB-2}_L = \left[ \frac{1}{L} \sum_{k \in [l_0, l_0+L]} \|\bar{\mathbf{e}}_2(k)\|_2^2 \right]^{\frac{1}{2}} \quad (7.34)$$

and the time-normalized mean square coordination error,

$$\text{MSE-COORD}_L = \left[ \frac{1}{L} \sum_{k \in [l_0, l_0+L]} \|\epsilon(k)\|_2^2 \right]^{\frac{1}{2}} \quad (7.35)$$

are used.

In both the optimal static and robust static control design considered here, we pick the stable closed-loop matrix,  $\mathbf{A}_c = \text{diag}(0.5, 0.5, 0.5, 0.5, 0.5, 0.5)$  and no sensing noise is assumed. First, consider the optimal static controller. The stabilization weighting matrices,  $\mathbf{W}_1$  and  $\mathbf{W}_2$  are picked such that greater emphasis are given for the lower density tones as the human visual system are more sensitive to the variations of these tones. We pick  $\mathbf{W}_1 = \bar{\mathbf{W}}_{e1}^T \bar{\mathbf{W}}_{e1}$  where

$$\bar{\mathbf{W}}_{e1}(i, j) = \begin{cases} w^3 & i = j, w = 1 - 0.2 \frac{(i-1)}{M_1-1} \\ 0 & i \neq j \end{cases} \quad (7.36)$$

Also  $\mathbf{W}_2 = \mathbf{W}_1$  is used. Finally  $\mathbf{R}_1$  and  $\mathbf{R}_2$  are arbitrarily set to be  $\mathbf{R}_1 = 1 \times 10^{-6} \mathbf{I}$ ,  $\mathbf{R}_2 = 1 \times 10^{-6} \mathbf{I}$ .

With no uncertainty (i.e.  $\delta_1 = 0, \delta_2 = 0$ ),  $\text{MSE-COORD}_L$  using  $l_0 = 500, L = 500$  are 0.1242 and 0.1030 for  $\mathbf{Y} = \mathbf{0}$  (no coordination) and  $\mathbf{Y} = 10\mathbf{I}$  (with coordination) respectively. However, the optimal static control is not designed for condition where there is uncertainty in the identification of the plant model. Imposing uncertainty with setting of  $\delta_1 = \delta_2 = 0.1$ , the  $\text{MSE-COORD}_L$  deteriorates to 0.1601 even with coordination.

To address this the robust static control approach is used. We pick  $\bar{\mathbf{W}}_{e1}$  as given in (7.36) and  $\bar{\mathbf{W}}_{e2} = \bar{\mathbf{W}}_{e1}$ . From (7.21) and (7.22), we consider two different settings of  $\alpha$  i.e.  $\alpha = 0$  and  $\alpha = 0.4$ . Figure 7.1 shows the  $\text{MSE-STAB-1}_L$ ,  $\text{MSE-STAB-2}_L$  and  $\text{MSE-COORD}_L$  performance metrics with different levels of uncertainty given by  $\delta_1 = \delta_2$ . With no uncertainty, a small coordination error (i.e. small  $\text{MSE-COORD}_L$  value) can be maintained

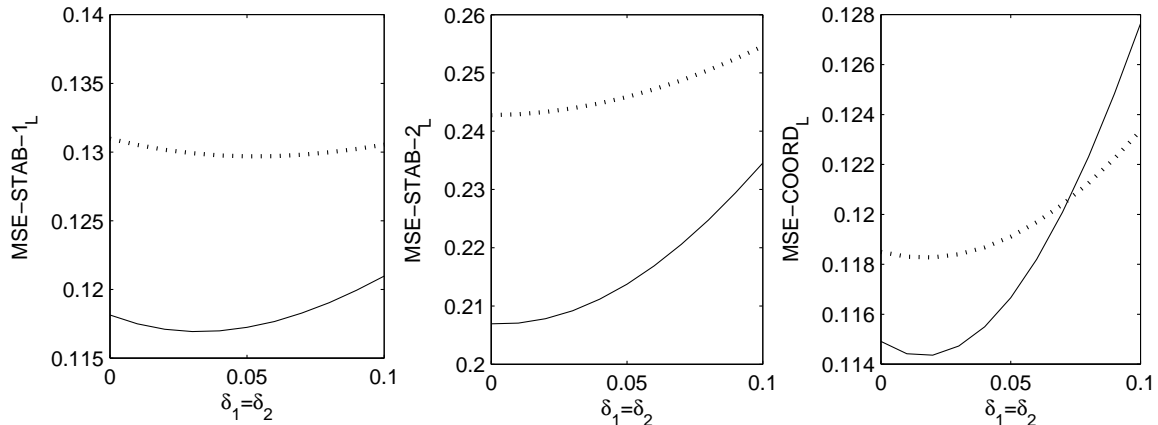


Figure 7.1: Stabilization and coordination performance using the robust static control approach of printer 1 and 2 with different uncertainty level given by different settings of  $\delta_1 = \delta_2$ .

using  $\alpha = 0.4$  at the expense of robustness. In this case, robustness can be regained with a loss in stabilization and coordination performance by setting  $\alpha$  to be small. Hence, appropriate settings of  $\bar{\mathbf{W}}_v$  and  $\bar{\mathbf{W}}_{ec}$  need to be selected to obtained the best tradeoff between robustness and performance.

### 7.2.1 Experiments

The proposed TRC stabilization and coordination system was experimentally tested using a Xerox Phaser 7700 xerographic printer as both printer 1 and printer 2. A X-rite DTP70 scanning spectrophotometer is used to measured the printed tone patches. Figure 7.2 shows the schematic of this setup. We do not have direct access to the xerographic actuators for both of the printers in this setup. To evaluate the proposed TRC stabilization and coordination control system, a virtual printer model is used to generate the response (tonal image) due to changes in the actuator inputs for both printers. The virtual printer model of both printers is given by (6.3) with no disturbance and no plant perturbation. By calibrating the printer such that it is an identity map at nominal condition, we can capture the effect of the disturbances on the TRC stabilization system. The disturbances at both printer are artificially introduced by injecting simulated sources  $\bar{\mathbf{d}}_{inject,1}(k)$  and  $\bar{\mathbf{d}}_{inject,2}(k)$  as shown in Figure 7.2.  $\bar{\mathbf{d}}_{inject,1}(k)$  and  $\bar{\mathbf{d}}_{inject,2}(k)$  are given by the form of equation (3.11) using similar parameters for the low and high-frequency bands as used in the simulation study. These disturbance models result in a visually significant color tone difference between the two printers, required for the purpose of this demonstration.

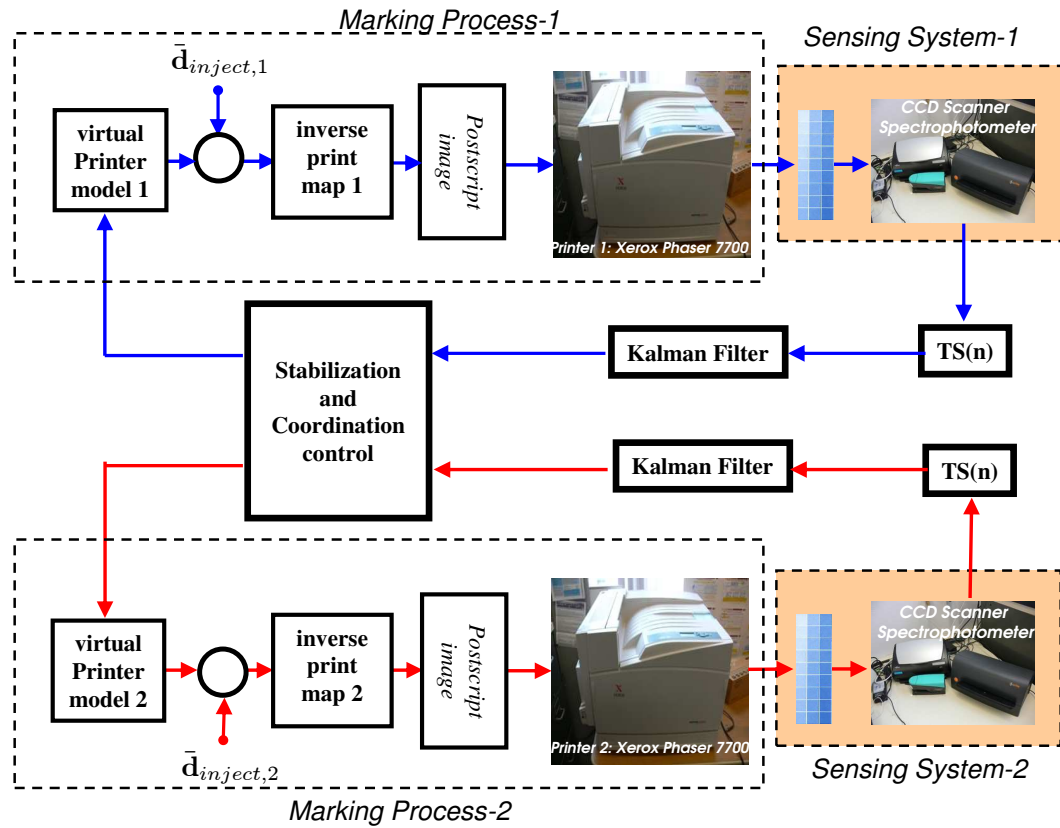


Figure 7.2: Experimental setup for sensing and control for stabilization and coordination of the TRCs

Figure 7.3 shows the TRC stabilization and coordination performance using the optimal static control approach for the all the CMY primary separations for three cases 1) no stabilization and coordination control, 2) stabilization with no coordination control and 3) with both stabilization and coordination control. Effective stabilization and coordination is demonstrated with case 3. This is shown in Figure 7.4 where the use of coordination and stabilization control resulted in printer-to-printer consistency as well as page-to-page consistency. Compared this result to the case where no coordination control is imposed as shown in Figure 7.5, the printer 1 to printer 2 color print mismatch is clearly visible.

### **7.3 Summary**

An optimal and robust static TRC stabilization and coordination control systems are proposed in this chapter that makes best use of the small number of xerographic actuators and limited sensing resources. Simulation and experimental results confirmed that good stabilization and coordination performance can be achieved with the proposed controllers. Given the uncertainty in the plant identification and the fact that manufactured units on the production line vary from unit to unit, the robust control approach can be used to gain the best tradeoff between coordination and stabilization performance and robustness.

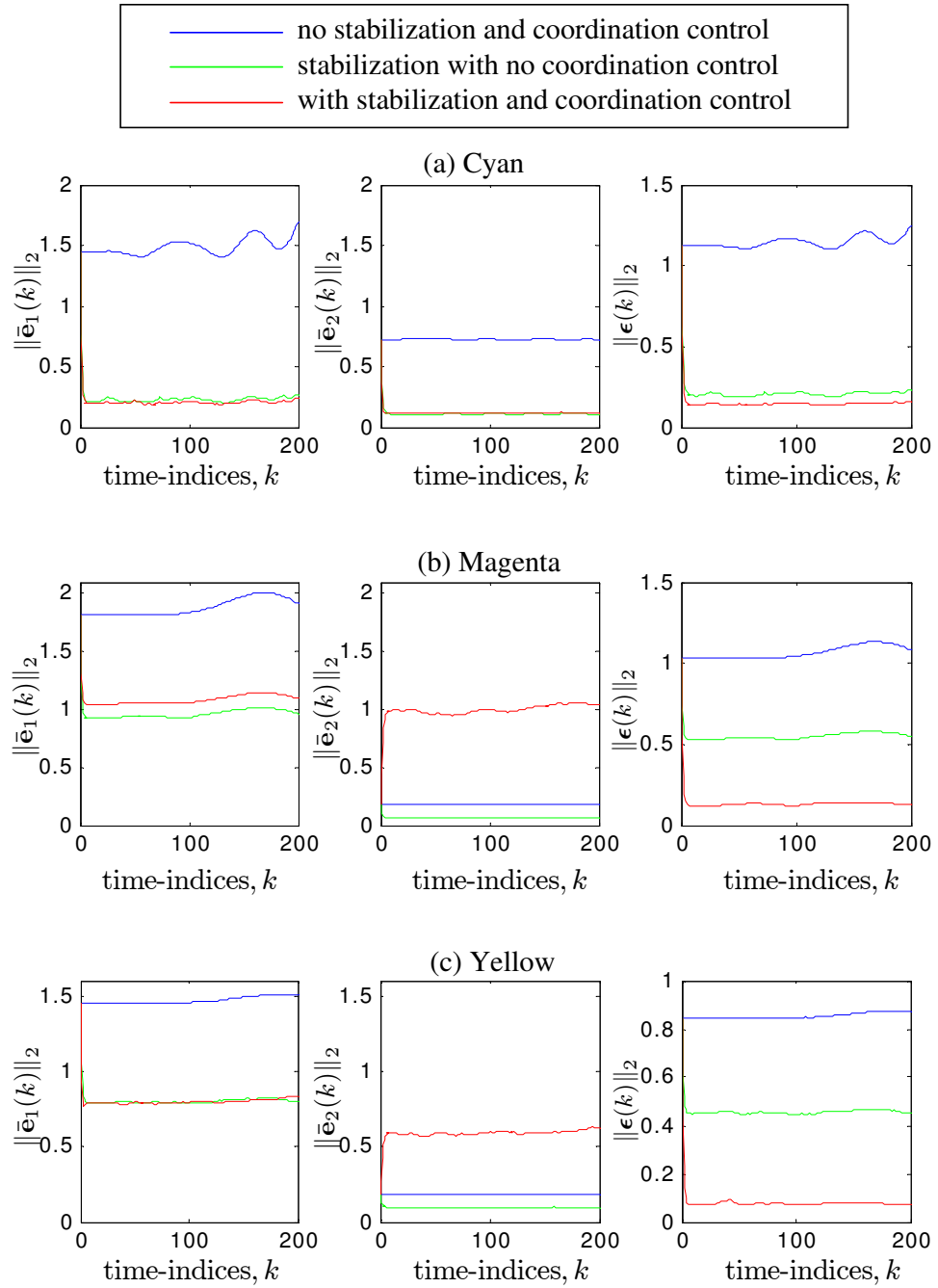


Figure 7.3: Experimental TRC stabilization and coordination performance for (a) cyan, (b) magenta and (c) yellow primary separations with optimal  $TS(1)$  sampling of printer 1 and printer 2. The optimal static control is used here with  $\mathbf{S} = 10\mathbf{I}$  for imposing coordination.

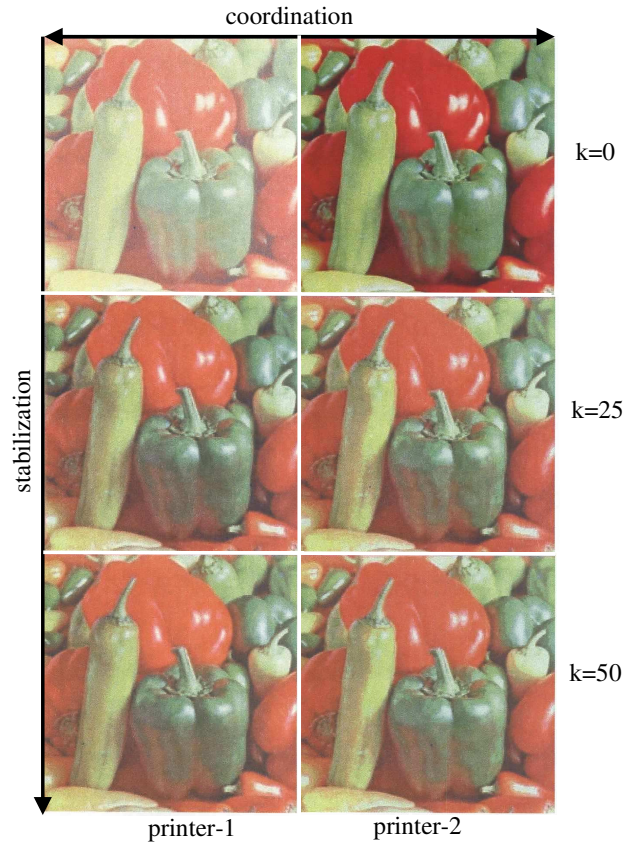


Figure 7.4: Color image printout of printer 1 (left) and printer 2 (right) at different time-step,  $k$  for the optimal static control approach *with* coordination. Note that color prints at  $k = 25$  and  $k = 50$  are much more closer to each other than with print at  $k = 0$  indicating stabilization of the color prints; and color prints at  $k = 50$  of both printers are closer than given at  $k = 0$  indicating coordination.

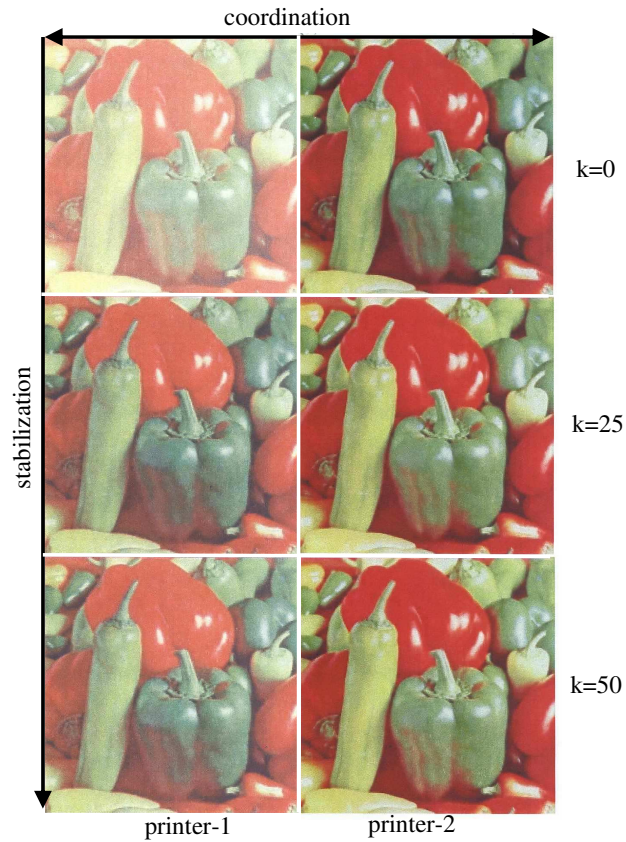


Figure 7.5: Color image printout of printer 1 (left) and printer 2 (right) at different time-step,  $k$  for the optimal static control approach *without* coordination. Note that, color prints at  $k = 25$  and  $k = 50$  are much more closer to each other than with print at  $k = 0$  indicating stabilization of the color prints; and color prints at  $k = 50$  of both printers are mismatch indicating a lack of coordination.

## Chapter 8

# Conclusion, Discussion and Future Work

The sensing and control for color consistency of a xerographic printing process has been discussed in this dissertation. Contributions have been made in the following areas:

1. Development and analysis of time-sequential sampling and causal reconstruction of  $1 + N$  dimensional temporal-spatial signal with a periodic Kalman filter (e.g. time-varying tone reproduction curve and time-varying color reproduction characteristics function)
2. Development of the algorithm to design the optimal time-sequential sampling for  $n = 1$  sample at each sampling instance (i.e. optimal  $TS(1)$  sampling) and extension to time-sequential sampling for  $n > 1$  samples at each sampling instance (i.e.  $TS(n), n > 1$  sampling). Validation of these design approaches show that they give better performance than other known sparse sampling approaches and perform as well as the complete sampling of the temporal-spatial signal.
3. Design, implementation and experimental validation of a tone/color stabilization control system that ensure page-to-page tone/color consistency despite disturbances on the print process.
4. Design, implementation and experimental validation of the tone coordination control system that ensures the printer-to-printer tone consistency despite the varying degree of disturbances on the different print systems.

These contributions are further detailed in the next section.



## 8.1 Contributions

### Contribution #1

Time-sequential sampling strategy is proposed here to increase the utility of the available feedback information. Time-sequential sampling was investigated in the 1980's and 1990's for video and time varying imaging applications (e.g. video and tomography). For time varying images, time-sequential sampling refers to sampling the image at different spatial locations at different sampling instances. Rastering in television is an illustration. The benefit is that by trading off temporal bandwidth with spatial bandwidth, the temporal bandwidth of the time varying image that can be captured faithfully beyond the Nyquist rate determined by the periodicity of the sampling scheme alone. In another perspective, the sampling rate can be reduced while retaining the same information content. In our context, time-sequential sampling means that at different sampling instances, different tones or different colors are sampled. This maximizes the information from the TRC/CRC samples and allows the time varying TRC or CRC to be captured (and subsequently reconstructed faithfully) even when only a small number of samples of the TRC/CRC are available at each sampling instant.

A periodic Kalman filter is proposed to causally reconstruct any arbitrary time-sequentially sampled TRC or CRC. Using Floquet theory, it is shown that time-sequential sampling and the subsequent reconstruction consists of the modulation, low pass and demodulation steps. Comparisons between two classical time-sequential sampling sequences (i.e. lexicographical and bit-reversed) show that the latter, which produces more uniform sampling and a more favorable spectral replication pattern, out-performs the former.

The general time-sequential sampling is demonstrated to be effective in monitoring the entire time-varying TRC even if only  $n = 1$  test patch is used. Tone consistency control performance is much better than if a small number (say  $n = 3$ ) of fixed tones are sampled, and is nearly as good as if the entire time varying TRC is available for feedback. These results were obtained using simple lexicographic and bit-reversed sampling sequences which are somewhat arbitrary. Indeed, none of these sampling sequences completely avoid aliasing, and neither the tonal resolution nor the sampling period was optimized. These deficiencies motivate the development of the optimal time-sequential sampling known here as optimal  $TS(1)$  and its extension to time-sequential sampling for arbitrary  $n > 1$  sampling points known here as the  $TS(n), n > 1$  sampling.

### Contribution #2

The design of the optimal  $TS(1)$  sampling is achieved in this dissertation through an optimization approach. Since the search space for the inter-sampling time,  $T$ , the number of discretization points,  $M_j$  at each  $j$ -th spatial dimension and the  $TS(1)$  sampling sequences could potentially be very large, finding the optimal solution is exceedingly complex and computationally prohibitive. To make the problem tractable, the  $TS(1)$  sampling sequences are restricted to be a  $1 + N$  dimensional lattice so that the lattice theoretic framework for time-sequential sampling in [13, 14] can be used. Although not all  $TS(1)$  sampling are lattices, we have shown that an optimal  $TS(1)$  lattice will be as good as an unconstrained optimal  $TS(1)$  sampling provided that the temporal-spatial signal is not undersampled temporally by more than a factor of 2. Moreover, since not all lattices form time-sequential sampling sequences, we obtain condition on the lattices that *do* form  $TS(1)$  sampling sequences. The space of such lattices is then parameterized by the inter-sampling time  $T$ , the number of spatial discretization at each dimension  $M_j$  and the spatial range  $A_j$ . The optimization problem of finding the time-sequential sampling lattice that avoids aliasing and maximizes the inter-sampling time can then be carried out with respect to these parameterizations. The lattice theoretic framework allows us to convert the temporal-spatial domain requirements into the frequency domain and vice versa. For example, the no-aliasing condition translates to a packing condition in the frequency domain, and the inter-sampling time maximization problem translates to one of minimizing the volume of a unit cell of the polar lattice in the frequency domain. Because the space of feasible non-aliasing  $TS(1)$  sampling lattices is compact (i.e. finite cardinality), optimization procedures that guarantee to terminate can be proposed and has been developed. Although finite, the cardinality of the set of feasible non-aliasing  $TS(1)$  sampling lattices can be very large. This results in optimization process with high computational complexity. For signal with high dimension this is particularly problematic. Therefore it is important to restrict the set of non-aliasing  $TS(1)$  sampling lattices to reduce the computational complexity of the optimization process.

Finally, this dissertation proposed the  $TS(n)$  sampling lattice design for  $n > 1$ , constructed with the optimal  $TS(1)$  sampling lattice. Under a signal uniformity condition, the proposed  $TS(n)$  sampling design avoids aliasing and hence guarantee good reconstruction performance. While this design does not necessarily achieve the optimal inter-sampling time, it allows us to arrive at a compromise between sensors hardware and inter-sampling time,  $T$ . This is of significant practical importance.

### **Contribution #3**

By using the proposed time-sequential sampling approach and reconstruction, we can fully characterize the time-varying TRC/CRC. This information is critical in the formulation of

effective closed-loop process level control to ensure that the CRC is as close to the identity map as possible. Unlike the control objective for most processes which is to control or regulate the output of the process, the color control problem consists of maintaining the *process itself* to be constant and stable. The difference is because every customer image to be printed can contain *many* and *any* possible colors which the xerographic printer must reproduce correctly all at once. Moreover, xerographic printers are often used in an on-demand manner in which consecutive customer images are different.

Both the CRC and TRC control formulations pose significant problems for sensing and control. It is because the CRC and TRC, as mappings, are potentially infinite dimensional whereas only a small number actuators and sensors are available. Even when each color coordinate is modestly discretized into 16 steps, the color quality of  $16^3 = 4096$  desired colors need to be kept track of for the color control problem, and 16 tones must be kept track of for the TRC control problem for each separation.

The small number of available xerographic actuators ( $m \approx 3$ ) per color separation poses a challenge to the control problem even if full information is available. In this dissertation a two-stage control strategy is proposed. In **STAGE-I** control, a curve fitting approach based on linear quadratic control with integrator dynamics [23, 27], makes use of the xerographic actuators (i.e. laser power, corotron voltage and development voltages) to compensate for variations in the tones/colors. This approach allows the designer to specify  $q$  tones or colors ( $q < m$ ) which will be precisely regulated while allowing the TRC/CRC to stay close to the desired values at the other tones. Alternatively in this stage, the robust curve-fitting technique proposed in [3] can also be employed. This is followed by the **STAGE-II** control which uses the image processing actuator to compensate for residual variations of the previous control stage.

#### **Contribution #4**

Additionally, in tasks requiring printing large quantities of the same image or collection of images with multiple printers, maintaining printer-to-printer tones/colors consistency is as important as page-to-page consistency. This is because if the printers are left uncoupled (i.e. maintaining the tones/colors consistency from printers-to-printers is not considered), then prints from one printer to another may appear different although the page-to-page consistency is maintained. In this dissertation we proposed control approaches to coordinate the tone responses of two xerographic printers subject to differing levels of disturbances. While coordination control of electromechanical systems has been well developed and matured [21], to the best of our knowledge there are no reported works on applying this to the

coordination of the TRC of different print systems.

An optimal static control approach and a robust static control approach are proposed for coordinating and stabilizing the TRCs. Without loss of generality, consider the case of two printers. The main idea here is to ensure that the TRC of printer 1 and printer 2 are close to their nominal (for stabilization) and to each other (for coordination) in a least-square sense. However since the number of tones to stabilize and coordinate is potentially high, and there are only a small number of effective actuators, there may be a loss of stabilization performance with increase in coordination. In our case, coordination performance may sometime be more important than in stabilization (i.e. tracking the individual tones of each printer) because maintaining printer-to-printer prints consistency may be more desirable.

## 8.2 Future Work

Research summarized in this dissertation points toward several areas where future work can be directed.

1. The design of the optimal  $TS(1)$  sampling proposed in this dissertation assumes that the underlying temporal-spatial signal is well described by its Fourier basis functions. This may not be the most efficient for the time-varying TRC/CRC that we wish to reconstruct or for that matter, any practical signal of interest. This is because signals in general have *localized* features. For example, in our case, certain colors that require nearly saturated primary color tones will have smaller variations and therefore lower color-space bandwidths than others. Since sinusoidal basis used in Fourier analysis are not localized, time-sequential sampling strategies computed based on spectral analysis will not take advantage of the localized temporal and non-temporal bandwidths trade-offs. Hence, wavelet descriptors will be more effective [60]. Since wavelet basis are localized, the signal class to be reconstructed can be represented using wavelet coefficients to capture the a-priori knowledge of local features. Optimal sampling strategy based on wavelet representations will likely be more efficient.
2. The design of the  $TS(n), n > 1$  sampling proposed in this dissertation is sub-optimal in that the inter-sampling time,  $T$  is *not* optimized. It remains to be seen how to specify the  $TS(n)$  sampling condition similar to that of Theorem 4.3 for the case of the  $TS(1)$  sampling that can be used to parameterize the  $TS(n), n > 1$  sampling lattice for optimization. Resolving this would enable the same algorithmic approach of finding the optimal  $TS(1)$  sampling design to be applied for the case of finding the optimal  $TS(n)$  sampling design.

3. In this dissertation, time-sequential sampling has been separately applied for sensing the TRC i.e. a  $1 + N = 2$  dimensional temporal-spatial signal and CRC i.e. a  $1 + N = 4$  dimensional temporal-spatial signal with great success. However from the Yule Nielsen Spectral Neugebauer (YNSN) model as shown in Figure 2.5, a color print is made up of the overlay of the Cyan, Magenta and Yellow (C, M, and Y) primaries and disturbances enter the printing process through 1) each of the primary TRC and 2) in the overlaying/mixing of the primaries. Thus if we were to sense only the primary TRCs, we will not account for the color mixing disturbances. Sensing the entire CRC map accounts for both disturbances but it is inefficient because the spectral contribution of the primaries enlarge the overall spectral content of the CRC. It remains to be seen how we can sense both the primary TRCs and CRC together, the former sense the variations due to the disturbances on the primary TRCs and the latter sense the variations due to the disturbances on the color mixing. Using this *hybrid* sampling approach, the efficiency of sensing the time-varying CRC could possibly be further improved.
4. The TRC and CRC stabilization and coordination control systems proposed in this dissertation have been evaluated through simulation and hardware-in-loop experimental results. The hardware-in-loop approach is employed here because we do not have direct access to the xerographic actuators. While these studies had been helpful in establishing the effectiveness of the proposed approach, actual implementation on a print system has not been achieved. Hence the effectiveness of such implementation has yet to be proven in the field.

# Bibliography

- [1] E.S.Hamby and D.M.Gross. A control-oriented survey of xerographic systems: basic concepts to new frontiers. In *American Control Conference ACC*, pages 2615–2629, Boston, June 2004. IEEE.
- [2] L.K.Mestha, M.Enzian, C.B.Duke, D.Platteter, S.B.Bolte, J.Lanphere, D.Viassolo, K.Mihalyov, R.Scarlata, and L.Purvis. Control elements in production printing and publishing systems: DocuColor iGen 3. In *Conference on Decision and Control CDC*, pages 4096–4108, Maui, December 2003. IEEE.
- [3] P.Y.Li and S.A.Dianat. Robust stabilization of tone reproduction curves for xerographic printing process. *IEEE Transactions on Control Systems Technology*, 9(2):407–415, March 2001.
- [4] L.K.Mestha, S.A.Dianat, and M.A.Scheuer. Device independent color controller and method. United States Patent: 6,236,474 B1, May 2001.
- [5] C.L.Chen and T.C.Chiu. Banding reduction in electrophotographic process. In *International Conference in Advanced Intelligent Mechatronics*, pages 81–86, Como, Italy, July 2001. IEEE/ASME.
- [6] D.N.Curry. Apparatus and method for providing accurate layer registration for multiple layer electronic printing system. United States Patent: 5,264,943, November 1993.
- [7] L.B.Schein. *Electrophotography and Development Physics*. Springer-Verlag, 1992.
- [8] L.K.Mestha and E.S.Saber. Systems and methods for sensing marking substrate area coverage using a spectrophotometer. United States Patent 7,110,142 B2, September 2006.
- [9] R.E.Groff, D.E.Koditschek, and P.P.Khargonekar. Invertible piecewise linear approximations for color reproduction. In *IEEE International Conference on Control Application*, pages 716–720, Trieste, Italy, September 1998. IEEE.

- [10] R.E.Groff, D.E.Koditschek, and P.P.Khargonekar. Representation of color space transformations for effective calibration and control. In *Non-Impact Printing 16*, pages 255–260, Vancouver, Canada, October 2000. IS&T.
- [11] J.P.Allebach. Analysis of sampling-pattern dependence in time sequential sampling of spatiotemporal signals. *Journal of Optical Society of America*, 71(1):99–105, January 1981.
- [12] J.P.Allebach. Design of antialiasing patterns for time-sequential sampling of spatiotemporal signals. *IEEE Transactions on Acoustic, Speech, and Signal Processing*, 32(1):137–144, February 1984.
- [13] N.P.Willis and Y.Bresler. Lattice-theoretic analysis of time-sequential sampling of spatiotemporal signals - part 1. *IEEE Transactions on Information Theory*, 43(1):190–207, January 1997.
- [14] N.P.Willis and Y.Bresler. Lattice-theoretic analysis of time-sequential sampling of spatiotemporal signals - part 2: Large space-bandwidth product asymptotics. *IEEE Transactions on Information Theory*, 43(1):208–220, January 1997.
- [15] M.A.Rahgozar and J.P.Allebach. Motion estimation based on time-sequentially sampled imagery. *IEEE Transactions on Image Processing*, 4(1):48–65, January 1995.
- [16] N.P.Willis and Y.Bresler. Optimal scan for time-varying tomography II: Theoretical analysis and fundamental limitations. *IEEE Transactions on Image Processing*, 4(5):654–666, May 1995.
- [17] D.P.Petersen and D.Middleton. Sampling and reconstruction of wave-number-limited functions in n-dimensional euclidean spaces. *Information and Control*, 5(4):279–323, December 1962.
- [18] R.S.Berns. *Billmeyer and Saltzman's Principles of Color Technology*. John Wiley and Sons Inc., New York, 3 edition, 2000.
- [19] T.E.Thieret, T.A.Henderson, and M.A.Butler. Method and control system architecture for controlling tone reproduction in a printing device. United States Patent: 5,471,313, November 1995.
- [20] H. Hindi and L.S.Crawford. Control in modern printing systems: Modular reconfigurable media paths, color consistency, fuser process, and registration (tutorial session). In *American Control Conference ACC*, St. Louis, June 2009. IEEE.

- [21] M.Tomizuka, J.S.Hu, T.C.Chiu, and T.Kamano. Synchronization of two motion control axes under adaptive feedforward control. *Journal of Dynamic Systems, Measurement, and Control*, 114(2):196–203, June 1992.
- [22] P.Y.Li, T.P.Sim, and D.J.Lee. Time-sequential sampling and reconstruction of tone and color reproduction functions for xerographic printing. In *Proceedings of the 2004 American Control Conference*, pages 2630–2635, Boston, United States, June 2004. IEEE.
- [23] T.P.Sim, P.Y.Li, and D.J.Lee. Using time-sequential sampling to stabilize the color and tone reproduction functions of a xerographic printing process. *IEEE Transactions on Control Systems Technology*, 15(2):349–357, March 2007.
- [24] T.P.Sim and P.Y.Li. Optimal time sequential sampling of tone reproduction function. In *American Control Conference ACC*, pages 5728–5733, Minneapolis, United States, June 2006. IEEE.
- [25] T.P.Sim and P.Y.Li. Optimal time-sequential sampling of color reproduction characteristic function. In *Non-Impact Printing 22*, pages 585–591, Denver, November 2006. IS&T.
- [26] T.P.Sim and P.Y.Li. Design of time-sequential sampling of temporal-spatial signals. *IEEE Transactions on Information Theory (in preparation)*.
- [27] T.P.Sim and P.Y.Li. Stabilization of color/tone reproduction curves using time-sequential sampling. In *Non-Impact Printing 20*, pages 204–209, Salt Lake City, November 2004. IS&T.
- [28] T.P.Sim and P.Y.Li. Representation of primary color tone reproduction curves for their stabilization. In *Non-Impact Printing 20*, pages 198–203, Salt Lake City, November 2004. IS&T.
- [29] T.P.Sim and P.Y.Li. Two-stage process control with optimized time-sequential sampling for tone consistency of a xerographic printing process. *IEEE Transactions on Control Systems Technology (in preparation)*.
- [30] T.P.Sim and P.Y.Li. On coordination and stabilization of two printers. In *American Control Conference ACC*, pages 3337–3342, Seattle, United States, June 2008. IEEE.
- [31] T.P.Sim and P.Y.Li. Direct color consistency control for xerographic printing. In *Non-Impact Printing 25*, pages 678–681, Louisville, September 2009. IS&T.



- [32] T.P.Sim and P.Y.Li. Determination of significant color map. In *Non-Impact Printing 25*, pages 327–330, Louisville, September 2009. IS&T.
- [33] H.Kipphan. *Handbook of Print Media*. Springer, Berlin, 2001.
- [34] G.Sharma. *Digital Color Imaging Handbook*. CRC Press, Inc., Boca Raton, FL, US, 2002.
- [35] R.Balasubramanian. The use of spectral regression in modeling halftone color printer. In *Optics and Imaging in Information Age*, pages 372–375, Rochester, NY, October 1996. IS&T/OSA.
- [36] M.Xia, E.Saber, G.Sharma, and A.M.Tekalp. End-to-end color printer calibration by total least squares regression. *IEEE Transactions on Image Processing*, 8(5):700–716, May 1999.
- [37] M.Rotea and C.Lana. A robust estimation algorithm for printer modeling. In *American Control Conference ACC*, pages 2636–2641, Boston, July 2004. IEEE.
- [38] J.P.Allebach. Time-sequential sampling in  $n$ -dimensional euclidean space. In *Proc. 21st Annual Allerton Conf. on Communications, Control, and Computing*, pages 110–119, Monticello, IL, October 1983.
- [39] A.E.Bryson. *Applied Linear Optimal Control*. Cambridge University Press, Cambridge, 2002.
- [40] M.L.Tyler and M.Morari. Estimation of cross directional properties: scanning versus stationary sensors. *AIChE Journal*, 41(4):846–854, June 1995.
- [41] S.Bittanti, P.Colaneri, and G.de Nicolao. The difference periodic Riccati equation for the periodic prediction problem. *IEEE Transactions on Automatic Control*, 33(8):706–712, August 1988.
- [42] J.J.Hench and A.J.Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Transactions on Automatic Control*, 39(6):1197–1210, June 1994.
- [43] N.Wereley. *Analysis and Control of Linear Periodically Time Varying Systems*. PhD thesis, Department of Aeronautics Astronautics, Massachusetts Institute of Technology, 1991.
- [44] N.Wereley and S.Hall. Frequency response of linear time periodic systems. In *Conference on Decision and Control CDC*, volume 6, pages 3650–3655. IEEE, December 1990.

- [45] H.Sandberg, E.Mollerstedt, and B.Bernhardson. Frequency-domain analysis of linear time-periodic systems. *IEEE Transactions on Automatic Control*, 50(12):1971–1983, December 2005.
- [46] M.Araki, Y.Ito, and T.Hagiwara. Frequency response of sampled data systems. *Automatica*, 32(4):493–497, April 1996.
- [47] F.Callier and C.A.Desoer. *Linear System Theory*. Springer Verlag, Berlin, 1991.
- [48] P.P.Vaidyanathan. *Multirate system and filter banks*. Prentice Hall, Englewood Cliffs, NJ, 1993.
- [49] M.A.Rahgozar and J.P.Allebach. A general theory of time-sequential sampling. *Signal Processing*, 28(3):253–270, September 1992.
- [50] J.W.S.Cassels. *An Introduction to the Geometry of Numbers*. Springer-Verlag, Berlin, 1971.
- [51] R.M.Mersereau and T.C.Speake. The processing of periodically sampled multidimensional signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 31(1):188–194, February 1983.
- [52] E. Dubois. The sampling and reconstruction of time-varying imagery with application in video systems. *Proceedings of the IEEE*, 73(4):502–522, April 1985.
- [53] H.Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag, New York, NY, 1993.
- [54] J.H.Conway and N.J.A.Sloane. *Sphere Packings, Lattices and Groups*. Springer, New York, NY, 1998.
- [55] X-Rite. vipPAQ. [http://www.xrite.com/product\\_overview.aspx?ID=802](http://www.xrite.com/product_overview.aspx?ID=802), 2009.
- [56] R.Smith and A.Packard. Optimal control of perturbed linear static systems. *IEEE Transactions on Automatic Control*, 41(4):579–584, April 1996.
- [57] M.J.Nadenau, J.Riechel, and M.Kunt. Wavelet-based color image compression: exploiting the contrast sensitivity function. *IEEE Transactions on Image Processing*, 12(1):58 – 70, May 2001.
- [58] W.Zeng, S.Daly, and S.Le. An overview of the visual optimization tools in JPEG 2000. *Signal Processing: Image Communication Journal*, 17(1):84 – 104, October 2001.

- [59] M.J.Nadenau. *Integration of Human Color Vision Models into High Quality Image Compression*. PhD thesis, École Polytechnique Federerale De Lausanne, 2000.
- [60] Stephane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, Cambridge, 1999.
- [61] E.Kreyszig. *Advanced Engineering Mathematics*. John Wiley & Sons, Singapore, 1993.
- [62] E.Agrell, T. Ericksson, A. Vardy, and K. Zeger. Closest point search in lattices. *IEEE Transactions on Information Theory*, 48(8):2001–2214, August 2002.

# Appendix A

## Colorspace Formulation

### A.1 Basic Color Science Formulation

Some basic color science terminologies are in order to properly describe this characterization. They are given as follows:

1. The spectral reflectance  $\mathbf{r}(\lambda)$  gives the electromagnetic radiation of reflected light on a color sample as a function of its wavelength  $\lambda$ . To quantify color, we need only to consider the visible wavelength taken here to be in range of  $\lambda_{min} = 380nm$  to  $\lambda_{max} = 730nm$ .
2. The visible spectral reflectance can be specified in terms of the tristimulus values  $XYZ$  by taking into account a standard illuminant,  $S(\lambda)$  and observer,  $[\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)]$ . The CIE D65 standard illuminant and 2<sup>o</sup> 1931 CIE standard observer are used in this dissertation. Then, the trimulus values,  $XYZ$  is given by:

$$XYZ = \frac{100\Delta\lambda}{\sum_{\lambda} S(\lambda)\bar{y}(\lambda)\Delta\lambda} \begin{bmatrix} S(\lambda_0)\bar{x}(\lambda_0) & \dots & S(\lambda_{N-1})\bar{x}(\lambda_{N-1}) \\ S(\lambda_0)\bar{y}(\lambda_0) & \dots & S(\lambda_{N-1})\bar{y}(\lambda_{N-1}) \\ S(\lambda_0)\bar{z}(\lambda_0) & \dots & S(\lambda_{N-1})\bar{z}(\lambda_{N-1}) \end{bmatrix} \mathbf{r}(\lambda) \quad (\text{A.1})$$

where  $\Delta\lambda$  is taken to be  $10nm$ , then  $N = 36$  with  $\mathbf{r}(\lambda) \in \mathbb{R}^{36 \times 1}$ .

3. The  $XYZ$  color representation is however not perceptually uniform. A nonlinear transformation is used to transform the  $XYZ$  tristimulus values to the perceptually

uniform  $L^*a^*b^*$  color representation as follows:

$$\begin{aligned} L^* &= 116f\left(\frac{Y}{Y_n}\right) - 16 \\ a^* &= 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right] \\ b^* &= 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right] \end{aligned} \quad (\text{A.2})$$

where  $f(x) = \begin{cases} x^{1/3} & x > 0.008856 \\ 7.787x + 16/116 & x \leq 0.008856 \end{cases}$  and  $X_n Y_n Z_n = (0.9504, 1, 1.0889)$  is the tristimuli of the white stimulus.

4. To compare the difference between two color samples given by their  $L^*a^*b^*$  values e.g.  $L^*a^*b_{s1}^*$  and  $L^*a^*b_{s2}^*$ , the CIE-94 color deference formulae,  $\Delta E_{94}^*$  is used.

$$\Delta E_{94}^* = \left[ \left( \frac{\Delta L^*}{k_L \cdot S_L} \right)^2 + \left( \frac{\Delta C_{ab}^*}{k_C \cdot S_C} \right)^2 + \left( \frac{\Delta H_{ab}^*}{k_H \cdot S_H} \right)^2 \right]^{1/2} \quad (\text{A.3})$$

where  $\Delta L^* := L_2^* - L_1^*$ ,  $\Delta a^* := a_2^* - a_1^*$ ,  $\Delta b^* := b_2^* - b_1^*$ ,  $\Delta C_{ab}^* = (a_2^{*2} + b_2^{*2})^{1/2} - (a_1^{*2} + b_1^{*2})^{1/2}$ ,  $\Delta H_{ab}^* = [\Delta a^{*2} + \Delta b^{*2} - \Delta C_{ab}^{*2}]^{1/2}$ , and  $S_L = 1$ ,  $S_C = 1 + 0.045C_{ab}^*$ ,  $S_H = 1 + 0.015C_{ab}^*$ ,  $C_{ab}^* = (a_1^{*2} + b_1^{*2})^{1/2}$  and  $k_L = k_C = k_H = 1$  at reference conditions.

5. To convert an image from the RGB format to the YCrCb format, the following linear transformation is used:

$$\begin{pmatrix} Y \\ Cr \\ Cb \end{pmatrix} = \begin{pmatrix} 0.2990 & 0.5870 & 0.1140 \\ 0.5000 & -0.41869 & 0.08131 \\ -0.16875 & -0.33126 & 0.50000 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (\text{A.4})$$

## A.2 Colorspace Parameterizations

In this section, a method of parameterizing the 3-dimensional to a 1-dimensional representation,  $f_{param} : \mathcal{C} \rightarrow \mathcal{T}$ ,  $\mathcal{C} \subset \mathbb{R}^3$ ,  $\mathcal{T} \subset \mathbb{R}$  given in [28]. The 3-dimensional representation can be re-parameterized to a 1-dimensional representation because the primary color curve is invariant to changes in the xerographic printing process due to disturbances or actuator inputs. Therefore all the primary color tones are well defined by a single color curve in the 3-dimensional color gamut. This fundamental property is the key in reducing the dimensionality of the 3-dimensional representation because it allows a fix parameterizations curve to be defined as a measure of the 3-dimensional representation.

By considering a single primary color, a color wedge of different tones ranging from input tones,  $x_1 = 0$  (no tone) to  $x_M = 1$  (solid tone) is printed and measured at *nominal* to obtained a set of experimental data in  $L^*a^*b^*$  colorspace. Let this be denoted by:

$$\Upsilon = \begin{pmatrix} \hat{L}^* \hat{a}^* \hat{b}^*(x_1) \\ \vdots \\ \hat{L}^* \hat{a}^* \hat{b}^*(x_M) \end{pmatrix} \in \mathbb{R}^{M \times 3} \quad (\text{A.5})$$

where  $\hat{L}^* \hat{a}^* \hat{b}^*(x_i)$  denotes the nominal measurement values of the input tone,  $x_i$  in the  $L^*a^*b^*$  colorspace. The parameterizations is in the form of a line where the concentration of the experimental data,  $\Upsilon$  is most significant (see Figure A.1). An effective procedure for performing this operation is the *Principal Component Analysis* (PCA).

To use Principal Component Analysis, consider the covariance matrix,  $\mathbf{S}^* \in \mathbb{R}^{3 \times 3}$  of  $\Upsilon$ . It is define by:

$$\mathbf{S}^* := E [(\Upsilon - E[\Upsilon])^T (\Upsilon - E[\Upsilon])] \in \mathbb{R}^{3 \times 3} \quad (\text{A.6})$$

The components of  $\mathbf{S}^*$ , denoted by  $s_{ij}^*$ , represent the covariances between the random variable components  $\hat{L}^* \hat{a}^* \hat{b}^*(x_i)$  and  $\hat{L}^* \hat{a}^* \hat{b}^*(x_j)$ . The component  $s_{ii}^*$  is the variance of the component  $\hat{L}^* \hat{a}^* \hat{b}^*(x_i)$  and indicates the spread of the component values around its mean value.

The gradient of the line where the concentration of the experimental data,  $\Upsilon$  is most significant is given by the vector direction of the principal component where the experimental data set has the most significant amounts of energy, denoted here by  $\mathbf{v}^* \in \mathbb{R}^3$ .  $\mathbf{v}^*$  is given by the direction of largest variance of the data that is  $\mathbf{v}^*$  is the eigenvector with the largest eigenvalue of  $\mathbf{S}^*$ .

The minimum and maximum values corresponding to the input tones,  $x_1 = 0$  and  $x_M = 1$  projected on this direction is given by:

$$\hat{\omega}(x_1) = \langle \mathbf{v}^*, \hat{L}^* \hat{a}^* \hat{b}^*(x_1) \rangle \quad (\text{A.7})$$

$$\hat{\omega}(x_M) = \langle \mathbf{v}^*, \hat{L}^* \hat{a}^* \hat{b}^*(x_M) \rangle \quad (\text{A.8})$$

where  $\langle \cdot \rangle$  gives the inner product of two vectors.

Hence the parameterizations function,  $f_{param} : \mathcal{C} \rightarrow \mathcal{T}$  is obtained by projection of the 3-dimensional measurement vector,  $L^*a^*b^*(x_i) \in \mathcal{C}$  in the  $L^*a^*b^*$  colorspace onto the principal

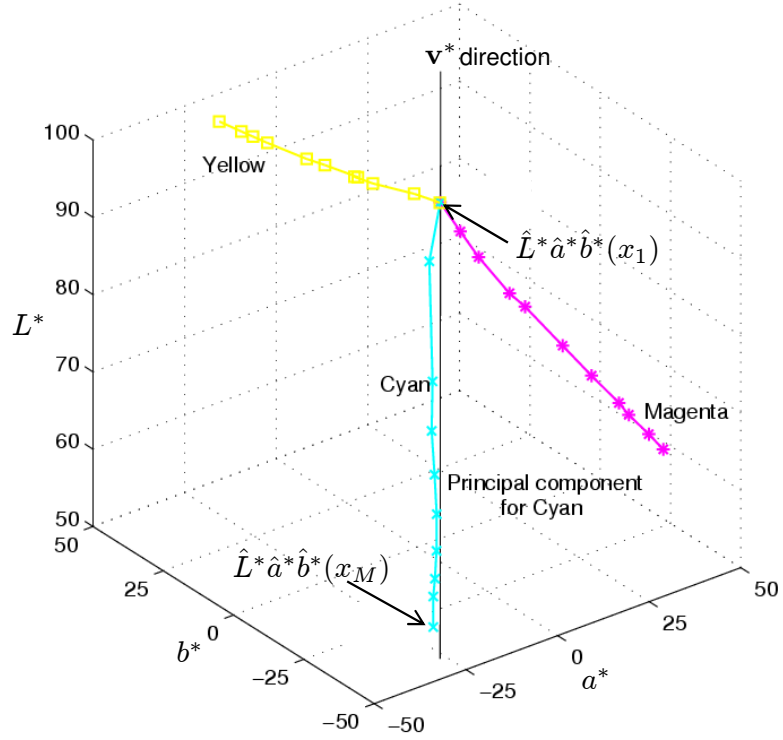


Figure A.1: Parameterizations by projection onto the principal component

component. By normalizing this projection such that the range of  $\bar{\omega}_i$ , we have:

$$\bar{\omega}_i = \frac{\langle \mathbf{v}^*, L^* a^* b^*(x_i) \rangle - \hat{\omega}(x_1)}{\hat{\omega}(x_M) - \hat{\omega}(x_1)}, \quad i = 1, \dots, M \quad (\text{A.9})$$

Figure A.1 shows the principal component and the associated parameters in the  $L^* a^* b^*$  colorspace for the primary color separation of cyan.

## Appendix B

# Formulation Characterizing Temporal-Spatial Frequency Content

### B.1 Matrix of Fourier Basis Function

To exhibit its tonal spectral contents, the small-magnitude, high-frequency TRC disturbance,  $\bar{\mathbf{d}}_h := [\bar{d}_h(1), \dots, \bar{d}_h(M)]^T \in \mathbb{R}^M$  is modeled by its discrete inverse Fourier transform (IDFT) i.e.

$$\bar{d}_h(p) = \frac{1}{M} \sum_{q=0}^{M-1} D(q) e^{i2\pi pq/M}, \quad p = 0, \dots, M-1$$

where  $D(q) \in \mathbb{C}$  is the complex Fourier coefficients, that is

$$D(q) = D_{real}(q) + iD_{imag}(q)$$

where  $D_{real}(q) \in \mathbb{R}, D_{imag}(q) \in \mathbb{R}$  are the real and imaginary coefficients respectively. For  $M$  odd, we have:

$$\bar{d}_h(p) = \frac{1}{M} \left\{ D_{real}(0) + \sum_{q=1}^{\frac{M-1}{2}} \left[ D(q) e^{i2\pi pq/M} + \bar{D}(q) e^{-i2\pi pq/M} \right] \right\}$$



where  $\bar{D}(q)$  is the complex conjugate to  $D(q)$ . Since  $D(q) + \bar{D}(q) = 2D_{real}$ , and by Euler formula,  $e^{i2\pi pq/M} = \cos\left(\frac{2\pi pq}{M}\right) + i \sin\left(\frac{2\pi pq}{M}\right)$  we have:

$$\bar{d}_h(p) = \frac{1}{M} \left\{ D_{real}(0) + 2 \sum_{q=1}^{\frac{M-1}{2}} \left[ D_{real}(q) \cos\left(\frac{2\pi pq}{M}\right) - D_{imag}(q) \sin\left(\frac{2\pi pq}{M}\right) \right] \right\} \quad (\text{B.1})$$

This can be rewritten in vector form as:

$$\bar{\mathbf{d}}_h = \mathbf{F} \cdot \mathbf{x}_h$$

where

$$\mathbf{x}_h = \left[ D_{real}(0), D_{real}(1), \dots, D_{real}\left(\frac{M-1}{2}\right), D_{imag}(1), \dots, D_{imag}\left(\frac{M-1}{2}\right) \right]^T$$

and

$\mathbf{F} =$

$$\frac{1}{M} \begin{pmatrix} 1 & 2 & \dots & 2 & 0 & \dots & 0 \\ 1 & 2 \cos\left(\frac{2\pi}{M}\right) & \dots & 2 \cos\left(\frac{\pi(M-1)}{M}\right) & -2 \sin\left(\frac{2\pi}{M}\right) & \dots & -2 \sin\left(\frac{\pi(M-1)}{M}\right) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 \cos\left(\frac{2\pi(M-1)}{M}\right) & \dots & 2 \cos\left(\frac{\pi(M-1)^2}{M}\right) & -2 \sin\left(\frac{2\pi(M-1)}{M}\right) & \dots & -2 \sin\left(\frac{\pi(M-1)^2}{M}\right) \end{pmatrix}$$

gives the Fourier basis matrix for  $M$  odd.

Similarly for  $M$  even, we have:

$$\bar{d}_h(p) = \frac{1}{M} \left\{ D_{real}(0) + D_{real}\left(\frac{M}{2}\right) \cos(\pi p) + 2 \sum_{q=1}^{\frac{M}{2}-1} \left[ D_{real}(q) \cos\left(\frac{2\pi pq}{M}\right) - D_{imag}(q) \sin\left(\frac{2\pi pq}{M}\right) \right] \right\} \quad (\text{B.2})$$

This can be rewritten in vector form as:

$$\bar{\mathbf{d}}_h = \mathbf{F} \cdot \mathbf{x}_h$$

where

$$\mathbf{x}_h = \left[ D_{real}(0), D_{real}(1), \dots, D_{real}\left(\frac{M}{2}\right), D_{imag}(1), \dots, D_{imag}\left(\frac{M}{2} - 1\right) \right]^T$$

and

$$\mathbf{F} = \frac{1}{M} \begin{pmatrix} 1 & 2 & \dots & 1 & 0 & \dots & 0 \\ 1 & 2 \cos\left(\frac{2\pi}{M}\right) & \dots & -1 & -2 \sin\left(\frac{2\pi}{M}\right) & \dots & -2 \sin\left(\frac{\pi(M-2)}{M}\right) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 \cos\left(\frac{2\pi(M-1)}{M}\right) & \dots & -1 & -2 \sin\left(\frac{2\pi(M-1)}{M}\right) & \dots & -2 \sin\left(\frac{\pi(M-1)(M-2)}{M}\right) \end{pmatrix}$$

gives the Fourier basis matrix for  $M$  even.

This basic derivation can be extended for the case of  $N = 3$ (CRC) given as follows:

$$\begin{aligned} \bar{d}_h(p_1, p_2, p_3) &= \frac{1}{M_1 M_2 M_3} \sum_{q_1=0}^{M_1-1} \sum_{q_2=0}^{M_2-1} \sum_{q_3=0}^{M_3-1} D(q_1, q_2, q_3) e^{i2\pi(p_1 q_1/M_1 + p_2 q_2/M_2 + p_3 q_3/M_3)} \\ &= \frac{1}{M_1 M_2 M_3} \left\{ \begin{aligned} &D_{real}(0, 0, 0) \\ &+ 2 \sum_{q_1=1}^{\frac{M_1-1}{2}} \begin{bmatrix} D_{real}(q_1, 0, 0) \cos\left(\frac{2\pi p_1 q_1}{M_1}\right) \\ - D_{imag}(q_1, 0, 0) \sin\left(\frac{2\pi p_1 q_1}{M_1}\right) \end{bmatrix} \\ &+ 2 \sum_{q_1=0}^{M_1-1} \sum_{q_2=0}^{\frac{M_2-1}{2}} \begin{bmatrix} D_{real}(q_1, q_2, 0) \cos\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_2 q_2}{M_2}\right) \\ - D_{imag}(q_1, q_2, 0) \sin\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_2 q_2}{M_2}\right) \end{bmatrix} \\ &+ 2 \sum_{q_2=0}^{\frac{M_2-1}{2}} \sum_{q_3=0}^{M_3-1} \begin{bmatrix} D_{real}(0, q_2, q_3) \cos\left(\frac{2\pi p_2 q_2}{M_2} + \frac{2\pi p_3 q_3}{M_3}\right) \\ - D_{imag}(0, q_2, q_3) \sin\left(\frac{2\pi p_2 q_2}{M_2} + \frac{2\pi p_3 q_3}{M_3}\right) \end{bmatrix} \\ &+ 2 \sum_{q_1=0}^{M_1-1} \sum_{q_3=0}^{\frac{M_3-1}{2}} \begin{bmatrix} D_{real}(q_1, 0, q_3) \cos\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_3 q_3}{M_3}\right) \\ - D_{imag}(q_1, 0, q_3) \sin\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_3 q_3}{M_3}\right) \end{bmatrix} \\ &+ 2 \sum_{q_1=0}^{M_1-1} \sum_{q_2=0}^{\frac{M_2-1}{2}} \sum_{q_3=0}^{\frac{M_3-1}{2}} \begin{bmatrix} D_{real}(q_1, q_2, q_3) \cos\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_2 q_2}{M_2} + \frac{2\pi p_3 q_3}{M_3}\right) \\ - D_{imag}(q_1, q_2, q_3) \sin\left(\frac{2\pi p_1 q_1}{M_1} + \frac{2\pi p_2 q_2}{M_2} + \frac{2\pi p_3 q_3}{M_3}\right) \end{bmatrix} \end{aligned} \right\} \end{aligned} \quad (\text{B.3})$$

## B.2 Finding the Ellipsoidal Compact Spectral Support, $\Theta(W, \mathbf{U})$

In this section, the ellipsoidal compact spectral support,  $\Theta(W, \mathbf{U})$  is deduced from a set of experimental time-varying data-set for each  $j$ -th output dimension collected from  $k = 0$  to  $k = K - 1$  i.e.:

$$\mathbf{\Gamma}_j = \begin{pmatrix} \omega_j(0, \mathbf{x}_1) & \omega_j(T_s, \mathbf{x}_1) & \dots & \omega_j(KT_s, \mathbf{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \omega_j(0, \mathbf{x}_{M_t}) & \omega_j(T_s, \mathbf{x}_{M_t}) & \dots & \omega_j(KT_s, \mathbf{x}_{M_t}) \end{pmatrix} \in \mathbb{R}^{M_t \times K} \quad (\text{B.4})$$

where  $T_s$  is the inter sampling time selected such that  $T_s < 1/2W_T$ .  $\mathbf{X}_s = \mathbf{A} \cdot / \mathbf{M}_s$  is the spatial resolution selected such that  $\mathbf{X}_s < 1 \cdot / (2\mathbf{W}_X)$ . Assume that  $(W_T, \mathbf{W}_X)$  is sufficiently large to accommodate all the temporal and spatial variations of interest. In this notation, for the case of the TRC,  $j = 1$  corresponding to the single tone output dimension. For the case of the CRC,  $j = 3$  corresponding to the  $L^*$ ,  $a^*$  and  $b^*$  output color dimension.

For each output dimension,  $j$  find the power spectral density of the data-set given by  $\mathbf{\Gamma}_j$  using the discrete Fourier transform analysis. Let the power spectral density be denoted by  $\Omega_d(f_l, \mathbf{u}_m)$  where  $f_l = l/KT_s$  for  $l \in [-K/2, K/2 - 1] \subset \mathbb{Z}$  and  $\mathbf{u}_m = \mathbf{m} \cdot / (\mathbf{M}_s \cdot * \mathbf{X}_s)$  for  $\mathbf{m} \in [-\mathbf{M}_s \cdot / 2, \mathbf{M}_s \cdot / 2 - 1] \subset \mathbb{Z}^N$ .

To find the region compact spectral support, consider the temporal-spatial frequencies points with significant frequency content. Let this be given by:

$$\mathcal{U} = \{(f_l, \mathbf{u}_m) : \Omega_d(f_l, \mathbf{u}_m) < \delta_t, \forall (l, \mathbf{m})\} \quad (\text{B.5})$$

where  $\delta_t > 0$  is a threshold value. The region of significant frequency content can then be given by the convex hull of  $\mathcal{U}$ ,  $H_{convex}(\mathcal{U})$ . Let  $\text{vol}[H_{convex}(\mathcal{U})]$  be the hyper-volume of  $H_{convex}(\mathcal{U})$  and  $\mathbf{r}_i$  specifies the  $i$ -th vertices position of  $H_{convex}(\mathcal{U})$ . Let the  $i$ -th intersection position of the ellipse with the  $\mathbf{r}_i$  vector be given by  $\mathbf{s}_i$  as shown in Figure B.1 for the case of  $1 + N = 2$ .  $\mathbf{s}_i$  can be computed by:

$$\mathbf{s}_i = \mathbf{r}_i / \|\mathbf{S}_\Theta \mathbf{r}_i\|_2 \quad (\text{B.6})$$

where  $\mathbf{S}_\Theta := \text{diag}(1/W, 1 \cdot / \mathbf{U})$ .

To find  $\Theta(W, \mathbf{U})$  we find an ellipse that tightly wrapped around  $H_{convex}(\mathcal{U})$ . This is achieved by minimizing the difference in the hyper-volume of the ellipsoid,  $\Theta(W, \mathbf{U})$  given by

$$\text{vol}[\Theta(W, \mathbf{U})] = \frac{\pi^{\frac{1+N}{2}}}{\left(\frac{N+3}{2}\right)} W U_t \quad (\text{B.7})$$

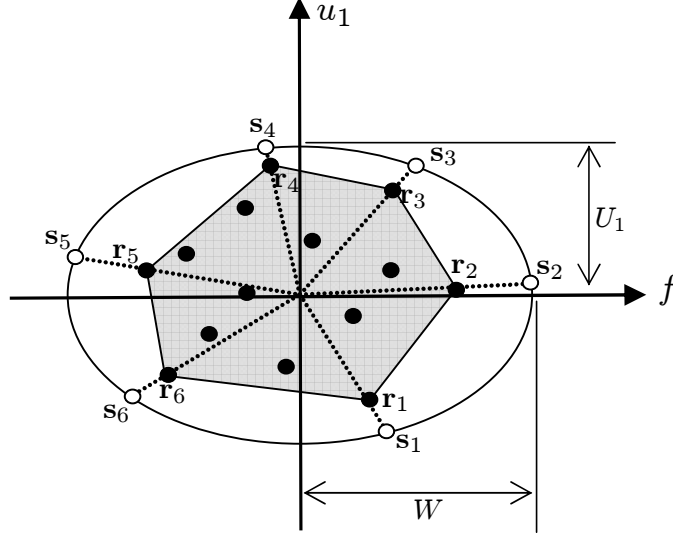


Figure B.1: The convex hull of the region of significant frequency content,  $H_{convex}(\mathcal{U})$  (shaded area), the ellipse,  $\Theta(W, U_1) \supseteq H_{convex}(\mathcal{U})$  and associated parameters.

where  $\Gamma(\frac{N+3}{2})$  is the Gamma function (see [61]) and  $U_t = \text{prod}(\mathbf{U})$ , and  $H_{convex}(\mathcal{U})$  subject to the constraint that  $\Theta(W, \mathbf{U}) \supseteq H_{convex}(\mathcal{U})$ . This is describe the the following problem statement:

$$\min_{(W, \mathbf{U})} \text{vol}[\Theta(W, \mathbf{U})] - \text{vol}[H_{convex}(\mathcal{U})] \quad (\text{B.8})$$

subject to  $\text{vol}[\Theta(W, \mathbf{U})] \geq \text{vol}[H_{convex}(\mathcal{U})]$  and  $\|\mathbf{s}_i\|_2 \geq \|\mathbf{r}_i\|_2, \forall i$ . This is a standard constrained nonlinear optimization problem that can be readily solved with `fmincon` command in Matlab.

## Appendix C

# Algorithms Used for Finding the Optimal $TS(1)$ Sampling

### C.1 Example of Optimal $TS(1)$ Sampling Design Proposed by Willis and Bresler

An example to design the optimal  $TS(1)$  sampling lattice proposed by Willis and Bresler [16] based on the lattice theoretic framework for time-sequential that they had proposed [13, 14] is described here. This example is taken directly from their *Example 1* [16] and is presented here to demonstrate their “by inspection” design approach. The spectral support considered in this paper, and in the example is given for the signal dimension of  $1 + N = 2$ , denoted here by  $\Theta$ .  $\Theta$  is shown for this example in Figure C.1(a). An important result used for realizing their proposed algorithm is restated here following notations used in this

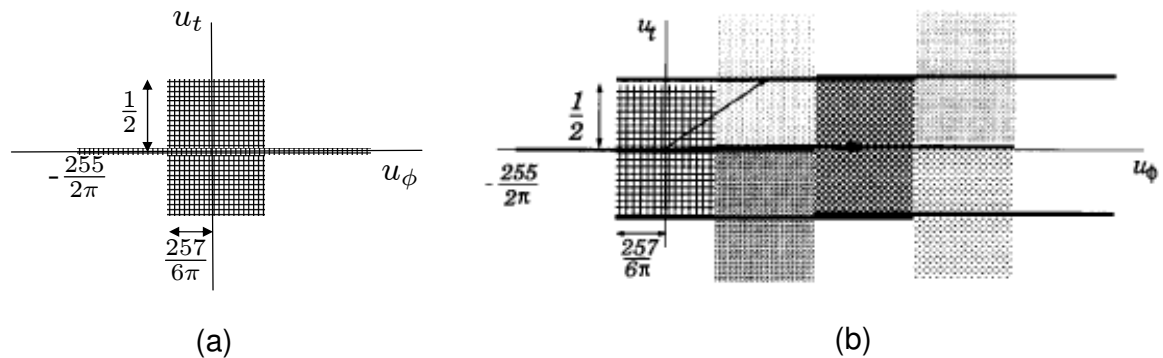


Figure C.1: (a) The assumed spectral support,  $\Theta$ ; (b) critical packing for spectral support,  $\Theta$  given by (a)

dissertation.

**Lemma C.1** *A rational lattice  $LAT(\mathbf{V}) \subset \mathbb{R}^2$  yields a  $TS(1)$  sampling lattice if and only if its polar lattice  $LAT(\mathbf{V}^*)$  has the basis:*

$$\mathbf{V}^* = \begin{pmatrix} \frac{p}{\pi} & \frac{l}{\pi} \\ q_1 & q_2 \end{pmatrix} \quad (\text{C.1})$$

with  $p$  and  $l$  coprime integers and  $q_1/q_2$  a rational number (i.e.  $\gcd(p, l) = 1$ , and  $q_1/q_2 \in \mathbb{Q}$ ).

The lemma enables the design of the optimal  $TS(1)$  sampling lattice as follows:

1. The spectral support  $\Theta$  for this example is shown in Figure C.1(a).
2. The critical packing for this spectral support shown in Figure C.1(b). The corresponding critical packing lattice has the basis:

$$\mathbf{V}^* = \begin{pmatrix} \frac{514}{3\pi} & \frac{257}{3\pi} \\ \epsilon & \frac{1}{2} + \epsilon \end{pmatrix} \quad (\text{C.2})$$

where  $\epsilon$  is an infinitesimally small number to allow the dc components of the spectrum to be as close as possible to each other without overlapping. This packing is 100% efficient and achieves the upper bound for the inter-sampling time,  $T$ , as  $\epsilon \rightarrow 0$ .

3. This packing is expanded slightly to the lattice with basis:

$$\mathbf{V}^* = \begin{pmatrix} \frac{172}{\pi} & \frac{87}{\pi} \\ \epsilon & \frac{1}{2} + \epsilon \end{pmatrix} \quad (\text{C.3})$$

which satisfies (C.1) for any rational  $\epsilon$ , and still packs  $\Theta$ . This lattice reach 99.6% of the upper bound for  $T$ . However, in practice, a finite value of  $\epsilon$  will have to be used, resulting in a correspondingly smaller  $T$ .

Note that this design process use the critical packing lattice given by (C.2) as guidance to find good packing of the  $\Theta$  “*by inspection*” that fulfills (C.1). The search heuristic consists of exploring lattices  $LAT(\mathbf{V}^*)$  “in the vicinity” of the starting point.

## C.2 Column-Reduced Hermite Normal Form Algorithm

The following gives an algorithm that constructs the Hermite Normal Form given in theorem 4.2, of an non-singular  $(1 + N) \times (1 + N)$  integer matrix,  $\mathbf{V}_z$ . The basic operation of the algorithm involves a row  $i$ , as well as columns  $s$  and  $t$  of the matrix  $\mathbf{V}_z$  with  $s < t$ . A sequence of elementary column operations (hence the column-reduced form) are performed so that  $v_{is} \leftarrow \gcd(v_{is}, v_{it}), v_{it} = 0$ .

**Theorem C.1** *Let  $\mathbf{V}_z = [\mathbf{v}_1, \dots, \mathbf{v}_{1+N}]$ ,  $\gcd(v_{is}, v_{it}) = r$ , and  $av_{is} + bv_{it} = g$ , where  $a$  and  $b$  are relatively prime. There exists an  $(1 + N) \times (1 + N)$  unimodular integer matrix  $\mathbf{Q}$  such that  $\mathbf{V}_z \mathbf{Q} = \bar{\mathbf{V}}_z$ , where*

$$\begin{aligned}\bar{\mathbf{v}}_l &= \mathbf{v}_l \text{ for } l \neq s, t \\ \bar{\mathbf{v}}_s &= a\mathbf{v}_s + b\mathbf{v}_t \\ \bar{\mathbf{v}}_t &= -\frac{v_{it}}{g}\mathbf{v}_s + \frac{v_{is}}{g}\mathbf{v}_t\end{aligned}$$

In particular,  $\bar{v}_{is} = g$  and  $\bar{v}_{it} = 0$ .

**Proof:** Take  $\mathbf{Q}$  to be an identity matrix in all but columns  $s$  and  $t$ . In column  $s$ , we have  $q_{ss} = a$ ,  $q_{ts} = b$  and  $q_{is} = 0$  otherwise. In column  $t$ , we have  $q_{st} = -v_{it}/g$ ,  $q_{tt} = v_{is}/g$ , and  $q_{it} = 0$  otherwise. It is readily verified that  $\mathbf{V}_z \mathbf{Q} = \bar{\mathbf{V}}_z$ , and  $\det(\mathbf{Q}) = av_{is}/g + bv_{it}/g = 1$ .

■

### Algorithm C.1 (The Hermite Normal Form algorithm)

*Initialization*  $i = 1$ .

**Step 1:** *Work on row  $i$ . Set  $j \leftarrow i + 1$ .*

**Step 2:** *Work on row  $i$  and  $j > i$ . If  $v_{ij} = 0$ , do nothing. Otherwise use the euclidean algorithm to find  $g = \gcd(v_{ii}, v_{ij})$  and  $a, b$  relatively prime such that  $av_{ii} + bv_{ij} = g$ . Set  $\mathbf{V}_z \leftarrow \mathbf{V}_z \mathbf{Q}$ , where  $\mathbf{Q}$  is the unimodular matrix described in theorem C.1, with  $s = i, t = j$ . If  $j < 1 + N$ , set  $j \leftarrow j + 1$  and return to **Step 2**. If  $j = 1 + N$ , go to **Step 3**.*

**Step 3:** *Work on row  $i$  and column  $i$ . If  $v_{ii} < 0$ , set  $\mathbf{V}_z \leftarrow \mathbf{V}_z \mathbf{Q}$ , where  $\mathbf{Q}$  multiplies column  $\mathbf{v}_i$  by  $-1$ .*

**Step 4:** Work on row  $i$  and column  $j < i$ , Set  $j \leftarrow 1$ . Set  $\mathbf{V}_z \leftarrow \mathbf{V}_z \mathbf{Q}$ , where  $\mathbf{Q}$  replaces column  $\mathbf{v}_j$  by  $\mathbf{v}_j - \lceil v_{ij}/v_{ii} \rceil \mathbf{v}_i$ . If  $j = i - 1$ , set  $i \leftarrow i + 1$ . If  $i > (1 + N)$ , stop. Otherwise return to **Step 1**. If  $j < i - 1$ , set  $j \leftarrow j + 1$  and return to **Step 4**.

### C.3 Closest Point Search Algorithm

The closest point search algorithm to determine  $\varphi_{min}$  for lattice  $LAT(\mathbf{H}_B^*)$  adapted from [62]. In [62] the algorithm searches for the closest lattice point for any given point in the multidimensional real space. Since the origin by default is a lattice point, using this algorithm to search for a lattice point closest to the origin will result in 0. Therefore, this algorithm has been modified to enable us to obtain  $\varphi_{min}$  as follows:

#### Algorithm C.2 (The closest point search algorithm)

**Step 1:** Compute  $\mathbf{H} = (\mathbf{H}_B^*)^{-T}$ .

**Step 2:** Let  $\mathbf{u} = \text{decode}(\mathbf{H})$

**Step 3:** Compute  $\varphi_{min} = \|\mathbf{H}_B^* \mathbf{u}^T\|_2$

The  $\text{decode}(\mathbf{H})$  algorithm where  $\mathbf{H} \in \mathbb{R}^{(1+N) \times (1+N)}$  is a lower triangular matrix i.e.

$$\mathbf{H} = \begin{pmatrix} h_{11} & 0 & \dots & 0 \\ h_{21} & h_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ h_{1+N,1} & h_{1+N,2} & \dots & h_{1+N,1+N} \end{pmatrix}$$

is given by the following pseudo-code. The operation  $\text{sgn}^*(z)$  returns -1 if  $z \leq 0$  and 1 if  $z > 0$ . Ties in the rounding operation  $\lfloor z \rfloor$  are broken arbitrarily.

- 1:  $bestdist \leftarrow \infty$  {current distance record}
- 2:  $k \leftarrow 1 + N$  {dimension of examined layer}
- 3:  $dist_k \leftarrow 0$  {distance to examined layer}
- 4:  $\mathbf{e}_k \leftarrow \mathbf{0}$
- 5:  $u_k \leftarrow \lfloor e_{kk} \rfloor$  {examined lattice point}
- 6:  $y \leftarrow (e_{kk} - u_k)/h_{kk}$
- 7:  $step_k \leftarrow \text{sgn}^*(y)$  {offset to the next layer}
- 8: **loop**
- 9:  $newdist = dist_k + y^2$
- 10: **if**  $newdist < bestdist$  **then**



```

11:   if  $k \neq 1$  then
12:     for  $i = 1$  to  $k - 1$  do
13:        $e_{k-1,i} = e_{ki} - yh_{ki}$ 
14:     end for
15:      $k \leftarrow k - 1$  {move down}
16:      $dist_k \leftarrow newdist$ 
17:      $u_k \leftarrow \lfloor e_{kk} \rfloor$  {closest layer}
18:      $y \leftarrow (e_{kk} - u_k)/h_{kk}$ 
19:      $step_k \leftarrow \text{sgn} * (y)$ 
20:   else
21:     if  $newdist \neq 0$  then
22:        $\hat{\mathbf{u}} \leftarrow \mathbf{u}$  {best lattice point so far}
23:        $bestdist \leftarrow newdist$  {update record}
24:        $k \leftarrow k + 1$  {move up}
25:     end if
26:      $u_k \leftarrow u_k + step_k$  {next layer}
27:      $y \leftarrow (e_{kk} - u_k)/h_{kk}$ 
28:      $step_k \leftarrow -step_k - \text{sgn} * (step_k)$ 
29:   end if
30: else
31:   if  $k = 1 + N$  then
32:     return  $\hat{\mathbf{u}}$ 
33:   else
34:      $k \leftarrow k + 1$  {move up}
35:      $u_k \leftarrow u_k + step_k$  {next layer}
36:      $y \leftarrow (e_{kk} - u_k)/h_{kk}$ 
37:      $step_k \leftarrow -step_k - \text{sgn} * (step_k)$ 
38:   end if
39: end if
40: end loop

```