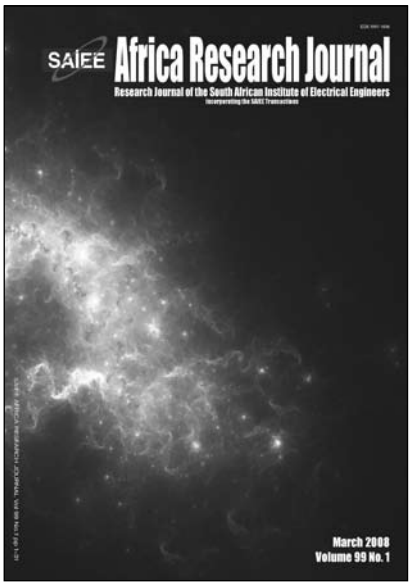
VOL 99 No 1
March 2008

SAIEE Africa Research Journal



GUEST EDITORIAL

IEEE AFRICON'07 : PART I

The 2007 IEEE Africon Conference was held in Windhoek, Namibia, on 26 – 28 September 2007.

This first of three special issues contains papers which were presented at the conference; the authors of which were subsequently invited to submit an expanded manuscript for publication. These papers cover a number of topics of considerable interest – in the general and specific sense.

Four papers are presented in this issue; three involve various aspects of wireless communications technology, while the last one covers systems design.

The first paper describes a graphical method for CDMA coding in mobile wireless systems. With bandwidth remaining a premium commodity, better coding methods are critical to make the best and most efficient use of that bandwidth.

The second paper also describes a method for signal coding in wireless communications. In this case, the focus is on using improved coding to reduce wasteful interference between and among wireless devices.

The third paper describes a CAD method of analysing on-chip spiral inductors in RF-CMOS circuit design. Inductors have always been an element of RF circuits, but they are not a natural part of integrated circuit process technology; their addition to process technology requires detailed analysis of the inductor and the parasitic elements inherent in an IC-based silicon manufacturing process.

The fourth paper evaluates the operating conditions for the sustained flight of solar-powered unmanned aerial vehicles (UAVs). UAVs have become a highly visible technology in recent years, most notably in their increasing size, range, capability – and potency. If a UAV can draw and store sufficient solar power to continue to operate during the dark part of the normal diurnal variation, then it can remain airborne indefinitely.

These papers provide an interesting sample of the work presented at IEEE Africon 2007; we hope that you find them interesting and informative.

Saurabh Sinha and Daniel Foty
Guest Editors

SPARSE GRAPH CODES ON A MULTI-DIMENSIONAL CDMA PLATFORM FOR APPLICATION IN MOBILE WIRELESS COMMUNICATION SYSTEMS

J.D. Vlok* and L.P. Linde**

** Defence, Peace, Safety and Security, Council for Scientific and Industrial Research, Meiring Naudé Road, Brummeria, Pretoria, South Africa*
*** Department of Electrical, Electronic and Computer Engineering, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa*

Abstract: This paper presents the uncoded and coded multiuser error performance results of a novel *super-orthogonal* four dimensional code division multiple access communication platform under additive white Gaussian noise and multipath fading channel conditions. The communication platform employs root-of-unity filtered constant envelope complex spreading sequences with zero cross-correlation properties. The uncoded communication platform employs multi-layered-modulation to improve spectral efficiency compared with existing multi-level modulation techniques, and the coded system uses the multiple dimensions to transmit redundancy to improve error performance. Three classes of sparse graph coding schemes are evaluated on the multi-dimensional communication platform. The channel codes include a three dimensional block-turbo-code with extended Reed-Muller constituent codes, low-density parity-check codes and repeat-accumulate codes.

Key words: Block Turbo codes (BTC), complex spreading sequences (CSS), channel modelling, log-likelihood ratio (LLR), low-density parity-check (LDPC) codes, multi-layered modulation (MLM), Multi-dimensional (MD), repeat-accumulate (RA) codes, sparse graph channel coding

1. INTRODUCTION

Multi-layered modulation (MLM) can be used to transmit parallel data streams with improved spectral efficiency compared with single-layer modulation, providing data throughput rates proportional to the number of modulation layers at performances equivalent to single-layer modulation. Alternatively, multiple layers can be used to transmit coded information to achieve improved performance at throughput rates equivalent to a single layer system.

An MLM four-dimensional (4D) code division multiple access (CDMA) building block employing super-orthogonal complex spreading sequences (SO-CSS) is used as MLM platform [1] for the evaluation of three classes of sparse graph codes. The communication platform including the transmitter, receiver and complex spreading sequences is discussed in Section 2. The encoding and decoding structures of the block-turbo-code and details regarding the construction of the constituent code's trellis structure and decoding algorithm are considered in Section 3. Section 4 briefly considers low-density parity-check (LDPC) and repeat-accumulate (RA) codes and Section 5 provides details regarding the channel simulator configuration. Finally Section 6 contains error performance results and Section 7 concludes the paper.

2. MODULATION PLATFORM

This section considers the basic 4D modulation platform [1] assuming perfect carrier synchronisation and spreading code lock.

2.1 Four Dimensional Transmitter

The basic 4D transmitter structure is essentially a Quadrature Phase Shift Keying (QPSK) modulator combined with spreading multipliers using the CSS $c(t) = c_r(t) + jc_i(t)$ as shown in Figure 1. The input sequence $d(t)$ is split up into four parallel data streams $d_1(t)$ to $d_4(t)$ by an serial-to-parallel (S/P) converter after which the upper two streams ($d_1(t)$ and $d_2(t)$) and the lower two streams ($d_3(t)$ and $d_4(t)$) are processed identically but separately. Two orthogonal spreading codes are used to spread each stream group as follows: data streams $d_1(t)$ and $d_2(t)$ are respectively multiplied by orthogonal spreading codes $c_r(t)$ and $c_i(t)$ to obtain $x_1(t)$ and $x_2(t)$ which are summed to obtain the in-phase component $y_1(t)$. The quadrature component $y_2(t)$ is obtained likewise from data streams $d_3(t)$ and $d_4(t)$. After modulating the in-phase and quadrature components onto quadrature carriers, the modulated components $z_1(t)$ and $z_2(t)$ are summed to produce the transmitted signal carrying 4 data bits per channel use.

$$\begin{aligned} s(t) &= \left[d_t(t)c_r(t) + d_2(t)c_i(t)\right]\cos(w_c t) \\ &\quad + \left[d_3(t)c_r(t) + d_4(t)c_i(t)\right]\sin(w_c t) \end{aligned} \tag{1}$$

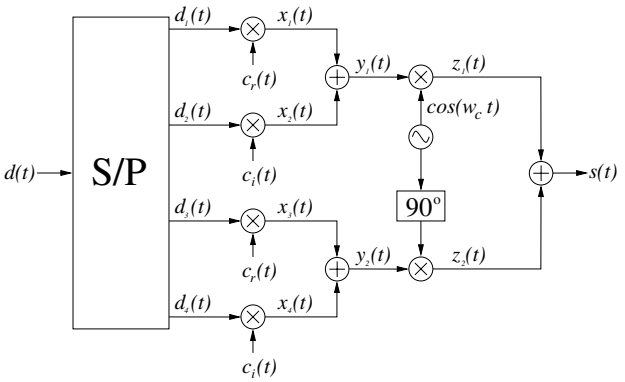


Figure 1: Four dimensional DSSS transmitter structure

2.2 Four Dimensional Receiver

The 4D receiver is a correlation-type demodulator as shown in Figure 2. The received signal *r(t)* is demodulated and despread to obtain approximations $\hat{d}_1(t)$ to $\hat{d}_4(t)$ of the original data streams. The use of unique families of CSSs contributes to the significantly simplified MUI-free receiver structure.

2.3 Complex Spreading Sequences

The performance of the CDMA system presented in this section depends primarily on the correlation characteristics of the CSSs used (see [2] for a thorough correlation study). The base vectors of the multi-dimensional (MD) space are the CSSs and it is therefore important to employ CSSs with excellent cross correlation characteristics between the real and imaginary components. One class of CSSs fulfilling this condition is Zero Cross Correlation (ZCC) [3] Super-Orthogonal (SO; i.e. zero-cross correlation for all time shifts) Constant-Envelope (CE) CSSs derived from Park-Park-Song-Suehiro (PS) [4] CSSs using a non-linear root-of-unity (RU) filtering technique [5]. The CE characteristic of ZCC CSSs provides the output signal *s(t)* of the 4D modulator with near-constant instantaneous and constant average output power by virtue of the orthogonality of the basis functions. The primary advantage of the power efficiency is the modulator's ability to operate close to the 1 dB saturation point of the power amplifier of the transmitter improving the communication range and/or handset battery life.

3. BLOCK TURBO CODE

3.1 Encoding Structure

A 3D block turbo code (BTC) [6] with *(N,K) = (2K,K)* constituent block codes is used to form a $(4K^3,K^3)$ encoder structure shown in Figure 3, by mapping the code blocks $\bar{c}_1$ to $\bar{c}_4$ directly to the 4D transmitter inputs $d_1(t)$ to $d_4(t)$ by removing the S/P converter (see Figure

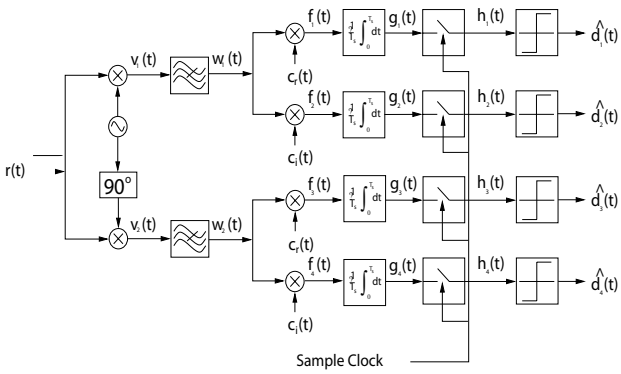


Figure 2: Four dimensional DSSS receiver structure

1). A 3D data block $\bar{x}$ is encoded by systematic row, column and depth encoders (1, 2 and 3 respectively) using row-to-column and row-to-depth block interleavers $\Pi_{1,2}$ and $\Pi_{1,3}$. The data and redundant information are transmitted simultaneously as a 4D signal $\bar{s} = [\bar{c}_0, \bar{c}_1, \bar{c}_2, \bar{c}_3]$ using SO-CSSs.

3.2 Decoding Structure

Correlation type demodulation (see Figure 2) is performed to calculate the soft value representing the kth; k = 1, . . . ,K^3, bit of the nth; *n = 0, 1, 2, 3*, code block $\tilde{c}_n$ using the equation:

$$\tilde{c}_{n,k} = \sum_{l=1}^{L.N_s} \alpha_{k,l} r_{k,l} \psi_{n,l} \qquad (2)$$

With $\alpha_{k,l}$ and $r_{k,l}$ respectively the lth instantaneous fading amplitude and received samples of the composite 4D signal containing the kth group of parallel codebits. $\psi_{n,l}$ is the lth sample of the orthonormal CSS base vector, distinguishing the nth dimension of the 4D signal [1]. *L* is the CSS length and N_s the number of samples per bit. The 4D demodulator output $\bar{y} = [\tilde{c}_0, \tilde{c}_1, \tilde{c}_2, \tilde{c}_3]$ is iteratively decoded using 3 interconnected soft-input soft-output (SISO) decoding modules [6] as shown in Figure 4.

The output of SISO module *m* is a 3D cube $\Lambda_{E,m}$; *m = 1, 2, 3*, containing the extrinsic log-likelihood ratio (LLR) of each data bit x_k, computed using the equation [7]:

$$\Lambda_{E,m}(x_k) = \Lambda(x_k \mid r_m) - \Lambda_{A,m}(x_k) - \Lambda_c(r_m \mid x_k) \qquad (3)$$

with $\Lambda_{A,m}(x_k)$ the prior LLR of x_k and $\Lambda_c(r_m \mid x_k)$ the channel LLR at the input of SISO module *m* and $\mathbf{r}_m$ is the received code block consisting of $\Pi_{1,m}\tilde{c}_0$ and $\tilde{c}_m$. Each block interleaver $\Pi_{m,m'}$ performs a 3D rotation on the input code block to translate the format processable by SISO module *m* to *m'*. Each SISO module computes the posterior LLR $\Lambda(x_k \mid r_m)$ employing a bidirectional

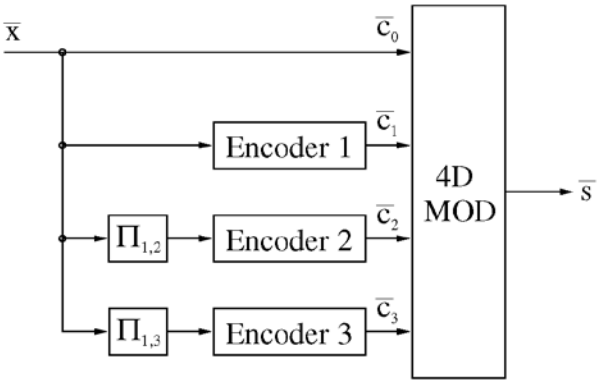


Figure 3: Encoder and 4D modulator structure

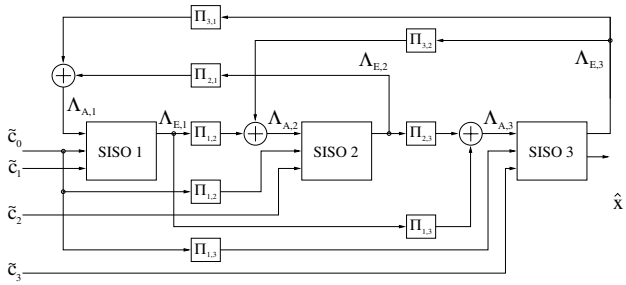


Figure 4: Three-dimensional block turbo decoder structure

Viterbi algorithm (VA) [8] with branch metric associated with x_k connecting states s and s' in the trellis:

$$BM_{s,s'}^{m,x_k} = \left(\tilde{c}_{n,k} - x_k\right)^2 - \frac{1}{2} x_k \Lambda_{A,m}(x_k) \qquad (4)$$

where n is determined by m, according to the structure shown in Figure 4. The estimated data block $\hat{x}$ is determined by performing hard decision quantisation on the posterior LLR of each data bit computed by SISO module 3.

3.3 Constituent Codes

Linear block codes namely binary Reed-Muller (RM) codes are used as constituent codes of the turbo coding scheme in this study. A soft-output bitwise decoding algorithm based on the trellis soft-output Viterbi algorithm (SOVA) is used to iteratively decode the BTC. Binary RM codes are chosen because of their relatively simple trellis structures.

Trellis Construction: The trellis structure of a linear block code can be constructed from its parity-check matrix [9]. The parity-check matrix of a linear *(N,K)* block code is given by:

$$H = \left[\bar{h}_1 \quad \bar{h}_2 \quad \cdots \quad \bar{h}_N\right] \qquad (5)$$

with $\bar{h}_i$ the ith column vector of H with the index $i = 1, 2, \ldots, N$ also denoting the ith trellis depth. Let $C_m = \left[c_{m,1}, \; c_{m,2}, \; \cdots, \; c_{m,N}\right]$ be the mth valid codeword with $m = 1, 2, \ldots, 2^K$ and let the states in the trellis be denoted as $\bar{S}_{m,i}$ in column vector format like $\bar{h}_i$. By including $\bar{S}_{m,0}$, the trellis can be constructed using the following recursion formula [9, 10]:

$$\bar{S}_{m,0} = \bar{0} \qquad (6)$$

$$\bar{S}_{m,i} = \bar{S}_{m,i-1} + c_{m,i}, \bar{h}_i \qquad (7)$$

Note that the trellis construction formula given in Equation 7 requires knowledge of all valid codewords, though all information needed to construct the trellis is inherently available in H. A trellis construction procedure presented in [10] starts at $S_{m,0} = 0$ and considers both possible binary values for each code bit $c_{m,i} \in [0,1]$ resulting in multiple paths, valid and invalid, in the trellis. The invalid paths are then removed by an expurgation procedure, i.e. removing all paths from the trellis that don't end in the zero state in the trellis at $i = N$. An alternative trellis expurgation technique is presented in this section.

The parity check matrix of a systematic $RM_{1,3}$ or RM(8, 4, 4) code is given as:

$$H = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (8)$$

from which the unexpurgated forward trellis structure from trellis depth $i = 0$ to $i = N$ is constructed. Note that Equations 6 and 7 can be rewritten to describe the trellis in reverse order as:

$$\bar{S}_{m,N} = \bar{0} \qquad (9)$$

$$\bar{S}_{m,i} = \bar{S}_{m,i+1} + c_{m,i+1}, \bar{h}_{i+1} \qquad (10)$$

Equation 10 follows from the fact that $c_{m,i}\bar{h}_i + c_{m,i}\bar{h}_i = \bar{0}$ in Equation 7. The unexpurgated reverse trellis structure is constructed by moving from trellis depth $i = N$ to $i = 0$ using Equations 9 and 10. By combining the unexpurgated forward and reverse trellises, the expurgated trellis (Figure 5) is obtained directly. The expurgated trellis structure (shown as solid branches in Figure 5) is the union of branches of the unexpurgated forward and reverse trellises (shown as dotted branches in Figure 5 not ending in the zero state at either trellis depths $i = 0$ or $i = N$).

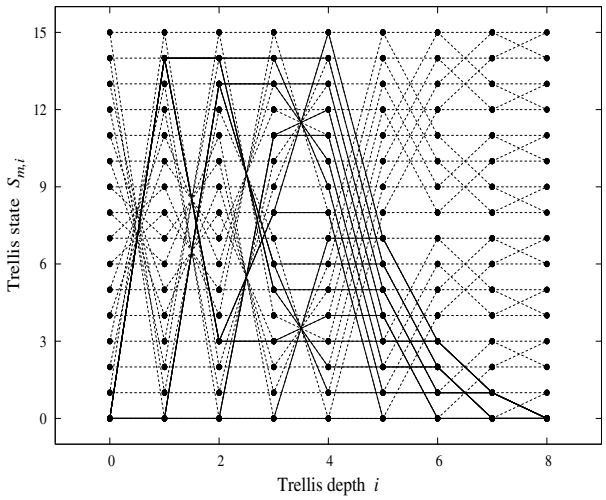


Figure 5: Unexpurgated (dotted branches) and expurgated (solid branches) trellis structure of the RM(8,4,4) code

Soft-Output Trellis Decoding: The reliability or posterior LLR of the transmitted information bits used in soft-output decoding algorithms can be expressed as [8, 11]:

$$\Lambda(x_k \mid r) = \ln\left[\frac{p(x_k = +1 \mid r)}{p(x_k = -1 \mid r)}\right] \quad (11)$$

$$= \ln\left(\sum_{l=1}^{L_t} e^{-PM_{l,k}^{x_k=+1}}\right) - \ln\left(\sum_{l=1}^{L_t} e^{-PM_{l,k}^{x_k=-1}}\right) \quad (12)$$

where L_t is the number of selected paths considered and $PM_{l,k}^{x=\pm1}$ is the lth path metric corresponding to $x_k = \pm1$. Equation 12 follows from the fact that $p(x_k = \pm1|r)$ is the sum of all probabilities of occurrences of $x_k = \pm1$ which correspond to the sum of probabilities of paths in the trellis with $x_k = \pm1$.

To calculate the probability $p(x_k = \pm1|r)$ exactly, all possible paths in the trellis corresponding to $x_k = \pm1$ must be considered and therefore $PM_{l,k}^{x=\pm1}$ corresponds to the path metrics of all possible paths (not only ML paths) in the trellis.

Using the max-log approximation [7, 11] Equation 12 can be simplified as:

$$\Lambda(x_k \mid r) \approx \min_l\left\{PM_{l,k}^{x=-1}\right\} - \min_l\left\{PM_{l,k}^{x=+1}\right\} \quad (13)$$

$$= PM_{ML,k}^{x=-1} - PM_{ML,k}^{x=+1} \quad (14)$$

where $PM_{ML,k}^{x=\pm1}$ is the path metric of the best candidate (ML) path with $x_k = \pm1$ in each trellis depth.

The soft-output decoding algorithm used in this study is a SOVA-based or bidirectional soft-output decoder [7, 8]. The algorithm calculates an estimate of the posterior LLR defined in Equation 11 by performing a forward-backward Viterbi algorithm on an expurgated trellis structure such as the *RM(8, 4, 4)* trellis shown in Figure 5. The algorithm computes and stores the path metrics in each trellis depth by running through the trellis independently in both forward and reverse directions. $\Lambda(x_k \mid r)$ is then calculated using the forward and reverse calculated path metrics and the branch metrics corresponding to $x_k = \pm1$ in each trellis depth.

The detailed forward-backward soft-output decoding algorithm used in this study is explained as follows:

- By running in the forward direction through the trellis according to Equation 7 starting at state $\overline{S}_{m,0}$ (see Equation 6), the forward path metrics are calculated as:

$$PM_{s,i}^{f} = \min_s\left(PM_{s',i-1}^{f} + BM_{s,s'}\right) \quad (15)$$

  where *(s, i)* indicates the trellis state $\overline{S}_{m,i}$ (as in Equations 6 and 7) in the ith trellis depth and (*s*?, *i*-1) indicates the trellis state in the previous trellis depth. $BM_{s,s?}$ is the branch metric of the branch connecting states *s* and *s*?. The path metric $PM_{s,i}^{f}$ is therefore the survivor path metric at state *s* up to trellis depth *i* in the forward direction. For each trellis coordinate *(s, i)* the survivor path metric should be stored.

- By running in the reverse direction through the trellis according to Equation 10 starting at state $\overline{S}_{m,N}$ (see Equation 9), the reverse path metrics are calculated as:

$$PM_{s,i}^{r} = \min_s\left(PM_{s',i+1}^{r} + BM_{s,s'}\right) \quad (16)$$

  where *(s, i)* indicates the trellis state $\overline{S}_{m,i}$ (as in Equations 9 and 10) in the ith trellis depth and (*s*?, *i*+1) indicates the trellis state in the next trellis depth. The path metric $PM_{s,i}^{r}$ is therefore the survivor path metric at state *s* up to trellis depth *i* in the reverse direction.

- For each trellis depth the best candidate path metric for $x_k = \pm1$ is calculated using the equation:

$$PM_{ML,k}^{x=\pm1} = \min_{x_k=\pm1}\left(PM_{s,i}^{f} + PM_{s',i+1}^{r} + BM_{s,s'}^{x_k=\pm1}\right) \quad (17)$$

where $BM_{s,s'}^{x_k=\pm 1}$ is the metric of the branch corresponding to $x_k = \pm 1$.

The posterior LLR $\Lambda(x_k \mid r)$ is finally calculated for each code bit using Equations 14 and 17.

The forward-backward decoding algorithm is used in each SISO decoding module in Figure 4 to calculate $\Lambda(x_k \mid r)$ with $k = 1, \ldots, N$ for each code bit in the received codeword. The extrinsic LLR $\Lambda_E(x_k)$ with $k = 1, \ldots, K$ is calculated for each information bit and passed to other SISO modules in the decoder using the posterior LLR $\Lambda(x_k \mid r)$ and Equation 3.

4. OTHER SPARSE GRAPH CODES

The two other classes of sparse graph channel coding schemes considered in this study are LDPC and RA codes, and are briefly considered in this section.

4.1 LDPC codes

Low-density parity-check (LDPC) codes [12] have received much research interest [13] and are powerful error-correcting codes outperforming turbo codes if the code length is large and the parity check matrix is designed correctly. In this study approximately regular LDPC codes constructed by random parity check matrix generation are used. Systematic sparse matrix encoding [14] is performed by writing the parity-check matrix as:

$$H = \begin{bmatrix} A & B \end{bmatrix} \tag{18}$$

with a non-singular sub-matrix **A** *(M × M)* and sub-matrix **B** *(M × K)*. The codeword can be written as $\bar{c} = [\bar{p} \quad \bar{x}]$ with $\bar{p}$ *(1×M)* the parity vector and $\bar{x}$ *(1×K)* the systematic information vector. From the code's null space and Equation 18 the following equation is obtained:

$$A\bar{p}^T + B\bar{x}^T = 0 \tag{19}$$

from which the parity vector can be calculated as:

$$\bar{p}^T = A^{-1}B\bar{x}^T \tag{20}$$

Decoding is performed using Pearl's belief propagation message-passing (MP) algorithm [15] on the code's factor graph. The algorithm takes as input the channel LLR $\Lambda_c(r_m \mid x_k)$ (see Equation 3) and provides as output an approximation of the posterior probability $p(t_n|r)$ for every transmitted bit.

4.2 RA codes

Repeat-Accumulate (RA) [15] codes are extremely simple Turbo-like [16] capacity-approaching codes, which allow linear-time encoding with performance results comparable to Turbo and LDPC codes. This study only considers the original non-systematic regular RA codes. RA codes can be decoded using a LLR MP algorithm on the code's factor graph, similar to decoding LDPC codes.

5. CHANNEL AND SIMULATION PLATFORM

AWGN and small-scale effects typically encountered in wide band mobile wireless communication systems, including frequency selective fading, are considered in this study. A channel model based on Clarke's model presented in [10, 17] (with a more accurate Doppler filter based on the design given in [18] - see [2] for a complete description of the channel simulator) is developed as part of the simulation platform to test and evaluate the performance of the CDMA system.

A multipath fading channel simulator with three statistically independent communication paths per user where each path has an unique Rician factor, Doppler spread and relative time delay similar to the model presented in [17] is used in the simulations. Each user's multipath channel has an exponentially decaying power delay profile as is typical in mobile wireless communications.

Shown in Figure 6 is the coded simulation platform. The data sequence *x(t)* originating from a random data generator, is encoded and then passed through an interfacing module to adapt the signal for transmission by the 4D transmitter. The code interface is an S/P module at the transmitter and a P/S module at the receiver. The channel encoder encodes a serial data sequence (or block) into a serial coded sequence and the 4D transmitter takes as input four parallel bit sequences given as four parallel estimated sequences at the output of the 4D receiver. It is therefore necessary to interface the encoder output to the transmitter input and receiver output to the decoder input. In the turbo coded system, the channel coding and interfacing modules are integrated into a single module by virtue of the turbo encoder and decoder structures. The channel state information (CSI, including the fading amplitude and phase) is extracted from the reference user's multipath fading channel simulator and used in the 4D demodulator. All other users' signals and AWGN are added to the reference user's signal before demodulation.

6. PERFORMANCE RESULTS

6.1 Block-Turbo Code

The uncoded and coded 4D communication systems are evaluated under AWGN and multipath fading conditions.

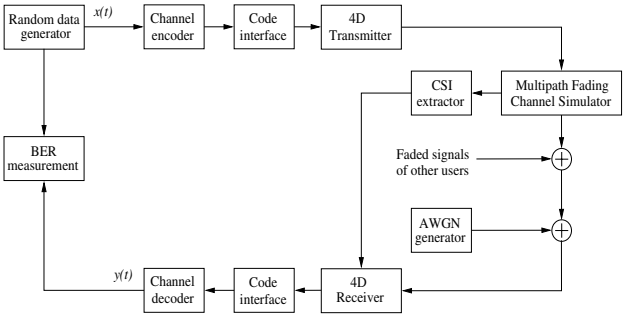


Figure 6: Coded simulation platform

Extended *RM* (12, 6) constituent codes are used to form a (864, 216) code with CSS parameters $L = 63$ and $N_s = 8$. Figures 7 and 8 show the AWGN and multipath fading error performance results respectively.

AWGN results of the BTC: Although the 4D uncoded system has data throughput equivalent to 16-QAM, the system maintains BPSK error performance under AWGN channel conditions, achieving a 4 dB gain relative to 16-QAM. The 3D BTC coded system has data throughput equivalent to BPSK and achieves a 6 dB coding gain over BPSK (about 5 dB from the Shannon limit) at P_e(bit) = 1×10^{-5}, after 5 decoding iterations. The results shown in Figure 7 are relevant for 1 to 5 users (the performance for 1 to 5 users lies on the same decoding iteration curve) with the number of users depending on the CSS length.

Multipath fading results of the BTC: Although a slight performance degradation (approximately 0.5 dB for E_b/N_0 larger than 8 dB) for 1 to 5 users is observed in multipath fading conditions for the uncoded system, practically identical multi-user results are achieved by employing the BTC. After 1 decoding iteration a 3 dB gain above the BPSK AWGN curve is evident and after 10 iterations an additional 2.5 dB is gained.

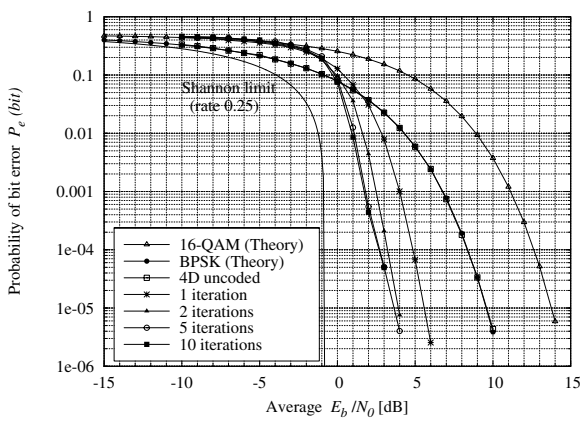


Figure 7: AWGN performance results for 1 to 5 CDMA users

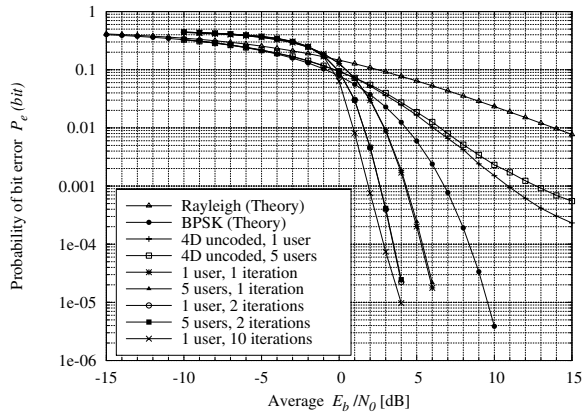


Figure 8: Multipath fading performance results for 1 to 5 CDMA users

6.2 Comparison of sparse graph codes

Comparing AWGN performance: Figure 9 shows the simulated AWGN error performance results of the uncoded 4D system, the 3D turbo code with the *RM* (8,4,4) constituent code (which is effectively a (256, 64) code) for 10 decoding iterations, the LDPC (256,64) code for 100 decoding iterations and the RA code for 100 decoding iterations. All curves shown are the single user error performance curves (which are identical to the performance curves for up to 5 CDMA users) using ZCC CSSs.

Comparing Multipath Fading Performance: Figure 10 shows the simulated multipath fading error performance results of the uncoded 4D system for 1 and 5 users, the 3D turbo code with the RM (8,4,4) constituent code for 1 and 5 users with 10 decoding iterations, the LDPC (256,64) code and the RA (256,64) code. The curves for the LDPC and RA codes shown in Figure 10 represent the performance curves for 1 to 5 users with 100 decoding iterations each.

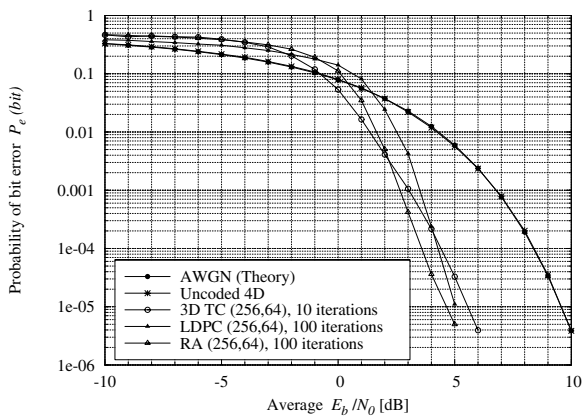


Figure 9: AWGN performance comparison of (256, 64) coding schemes

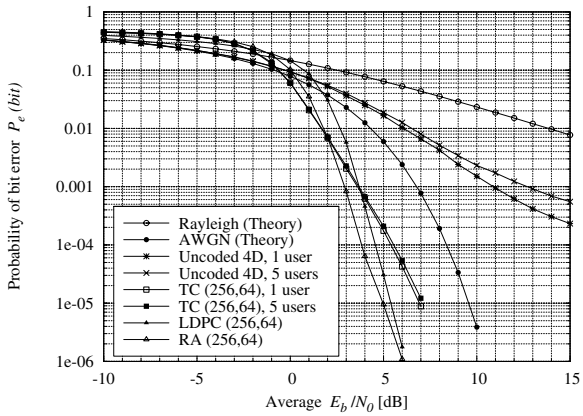


Figure 10: Multipath performance comparison of (256, 64) coding schemes

7. CONCLUSION

This paper presented the error performance results of the uncoded and coded 4D communication system presented in [1], showing BPSK error performance for all CDMA users under AWGN channel conditions. The three coding schemes were adapted to the 4D system and show comparable error performance results (within 2 dB), though at low E_b/N_0 values (around 0 dB) the turbo code performs a little better than the other two coding schemes and at 3 dB and higher the LDPC and RA codes outperform the turbo code. The RA code shows the best overall error performance. It is also clear that the turbo code requires less decoding iterations than the other coding schemes, though different decoding structures and algorithms (with different levels of complexity) are employed for each scheme. The 3D block-turbo-code (with RM(12,6) constituent codes) yielded an approximate 4 to 6 dB coding gain in AWGN and 3 to 6 dB coding gain in multipath fading channel conditions over BPSK at P_e(bit) = 10^{-5}, with decoding complexity of each constituent decoder equivalent to the bidirectional VA.

8. REFERENCES

[1] L. P. Linde and J. D. Vlok, "Power and spectrally efficient four dimensional super-orthogonal WCDMA building block for next generation wireless applications," *IEEE Communications Letters*, vol. 10, no. 7, pp. 519–521, July 2006.

[2] J. D. Vlok, "Sparse graph codes on a multi-dimensional WCDMA platform," Master's dissertation, University of Pretoria, South Africa, 2007.

[3] I. Pryra, L. P. Linde, and S. A. Swanepoel, "New family of constant envelope root-of-unity filtered complex spreading sequences with zero cross-correlation properties," *Proceedings of 6th AFRICON IEEE Conference*, vol. 1, pp. 299–304, October 2002.

[4] N. Suehiro, "A signal design without co-channel interference for approximately synchronized CDMA systems," *IEEE JSAC*, vol. 12, pp. 837–841, June 1994.

[5] L. P. Linde and M. Lötter, "Spread spectrum modulator and method," South African Complete Patent 96/0355, January 17, 1996, earliest priority claimed: ZA 94/9014, 14 November 1994.

[6] B. Yin and M. R. Soleymani, "Design and implementation of three dimensional block turbo codes," *Proceedings of 2003 IEEE Canadian Conference on Electrical and Computer Engineering - CCECE2003,* Montreal, Canada, pp. 1625–1628, May 2003.

[7] R. H. Morelos-Zaragoza, *The Art of Error Correcting Coding*, John Wiley & Sons Limited, Chichester, UK, first edition, 2004.

[8] W. Feng and B. Vucetic, "A list bidirectional soft output decoder of turbo codes," *Proceedings of the International Symposium on Turbo Codes and Related Topics*, pp. 288–292, September 1997.

[9] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, March 1974.

[10] L. Staphorst, "Viterbi decoded linear block codes for narrowband and wideband communication over mobile fading channels," Master's dissertation, University of Pretoria, South Africa, 2005.

[11] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo-Coding, Turbo Equalisation and Space-Time Coding for transmission over Fading Channels*, John Wiley & Sons Limited, Chichester, UK, first edition, 2002.

[12] R. G. Gallager, "Low-density parity-check codes," Ph.D. thesis, M.I.T., Cambridge, Massachusetts, 1963.

[13] T. J. Richardson and R. L. Urbanke, "The Renaissance of Gallager's Low-Density Parity-Check codes," *IEEE Communications Magazine*, vol. 41, no. 8, pp. 126–131, August 2003.

[14] R. M. Neal, "Sparse matrix methods and probabilistic inference algorithms - Part I: Faster encoding for low density parity check codes using sparse matrix methods," IMA Programs on Codes, Systems and Graphical Models, Tech. Rep., August 1999.

[15] D. J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, Cambridge, UK, first edition, 2005.

[16] D. Divsalar, H. Jin, and R. J. McEliece, "Coding theorems for Turbo-Like codes," *Proceedings of 36th Allerton Conference on Communication, Control, and Computing*, pp. 201–210, September 1998.

[17] L. Staphorst and L. P. Linde, "Evaluating Viterbi decoded Reed-Solomon block codes on a complex spreaded DS/SSMA CDMA system: Part II -

Channel model, evaluation and results," *Proceedings of AFRICON IEEE Conference*, pp. 335–340, 2004.

[18] C. Komninakis, "A fast and accurate Rayleigh fading simulator," *Proceedings of GLOBECOM Global Telecommunications IEEE Conference*, vol. 6, pp 3306–3310, December 2003.

COMBINED SPECTRAL SHAPING CODES AND OFDM MODULATION FOR NARROWBAND INTERFERNCE CHANNELS

K. Ouahada, H. C. Ferreira, A. J. H. Vinck*, A. J. Snyders and T. G. Swart

Department of Electric and Electronic Engineering Science, University of Johannesburg, South Africa
** Institute of Experimental Mathematics, University of Duisburg-Essen, Germany*

Abstract: Distance Preserving Mappings is a relatively new technique finding its way to a wide use in the coding field. New codes mapping convolutional code outputs onto spectral nulls codewords have the purpose of generating a code with a certain spectrum specification and error correction capabilities, which can be exploited by the Viterbi decoder. Combining this class of new codes with an Orthogonal Frequency Division Multiplexing modulation scheme is actually a tool to reduce Electromagnetic Compatibility Interference in selected subbands or to cancel the narrowband interference potentially caused by, amongst others, high frequency radio transmissions. Taking into consideration the fact that a narrowband noise source is almost similar to a modulated orthogonal frequency division multiplexing carrier, and also the periodicity of the nulls in our spectrum, we are able to cancel the narrowband signal and its sideband interferences without using traditional techniques of cancellation, which is based on filtering. Shaping our spectrum at the transmitter is a technique, which can be exploited at the receiver to suppress any narrowband signal interference, which is regarded as noise. We will present in this paper the technique of cancellation and the coding gain that can be obtained when comparing to the uncoded data.

Key Words: Distance-preserving mappings, Spectral nulls codes, Spectral shaping codes, OFDM modulation, Narrowband interference.

1. INTRODUCTION

We will describe in this paper the construction of a new code by using the Distance Preserving Mappings (DPM) technique [1], here used to obtain Spectral Nulls (SN) codes that, when transmitted serially, will have a spectrum with zeros occurring at $f = 1/kT$, where k is an integer or a ratio of relatively prime integers.

We furthermore investigate the performance of this code when combined with Orthogonal Frequency Division Multiplexing (OFDM) modulation, which has proven its ability to deal with broadband frequency selective fading channels in wire-line and wireless communication systems. The creation of spectral nulls at certain frequencies is to achieve narrowband interference cancellation in order to minimize the problem of Electromagnetic Compatibility Interference (EMC), while also exploiting the favourable error correction capabilities of the code with spectral nulls codes (SNC). We have to control the spectrum of our transmitted data in such a way that we do not cause interference to other spectrum users, e.g. high frequency (HF) Radio.

It is important to mention that the use of the Fast Fourier transform (FFT) in the OFDM modulation scheme will not allow us to switch off any of the affected subcarriers by the NBI, as a technique of avoiding the corruption of the data at that corresponding carrier, because of the strong link between the number of chosen carriers and the number of samples. This is why the FFT depends on the total number of subcarriers to achieve a very accurate transformation.

The paper is organized as follows: In Section 2 we present briefly the way of designing Spectral Nulls Codebooks. Section 3 introduces the technique of DPM as applied in the paper. The modified OFDM transmitter will be presented in Section 5. The cancellation and correction of the NBI results are in Section 6, and finally in Section 7 a conclusion of the performed analysis will be drawn.

2. SPECTRAL NULLS CODES DESIGN

The design of a baseband data stream with spectral nulls occurring at certain frequencies [2], is based on the consideration of the vector $x = (x_1, x_2, \ldots, x_M)$, $x_i \in \{-1, +1\}$ with $1 \leq i \leq M$, to be an element of a set S, which is called the codebook of codewords with elements in $\{-1, +1\}$, here we represent -1 as 0, and the application of the Fourier transform to those

codewords, to get:

$$X(w) = \sum_{i=1}^{M} x_i e^{-jiw}, -\pi \leq w \leq \pi. \quad (1)$$

Having nulls at certain frequencies is the same as having the power spectral density function equal to zero at those frequencies. This means that $H(w) = 0$, where

$$H(w) = \frac{1}{Mn} \sum_{i=0}^{M-1} |X^{(i)}(w)|^2. \quad (2)$$

It can be seen from Equation 2, that the power spectral density (PSD) depends on the frequency value. So we can design a codebook to generate nulls at certain chosen frequencies. Usually for simplification we present the codeword length, M, as an integer multiple of k, then

$$M = ks,$$

where $f = r/k$ represents the spectral nulls at rational sub multiples r/k [3]. We have to satisfy

$$A_1 = A_2 = \cdots = A_k, \quad (3)$$

where

$$A_i = \sum_{\lambda=0}^{s-1} x_{i+\lambda k}, \quad i = 1, 2, \ldots, k. \quad (4)$$

If all the codewords in a codebook satisfy these equations, the codebook will exhibit nulls at the required frequencies.

The spectral nulls codebook for $M = 6$ with $k = 3$ and $s = 2$ is:

$$C = \begin{Bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{Bmatrix}$$

The power spectral density is shown in Figure 1.

3. DISTANCE PRESERVING MAPPING TECHNIQUE

Distance Preserving Mappings is a coding technique which maps the outputs of a Conventional Convolutional Code (CCC) onto other codewords, which can be either binary or non-binary [4], from a code with lesser error-correction capabilities. The purpose behind this technique is to firstly obtain suitably constrained output code sequences and secondly to exploit the error correction characteristics of the new code with the use of the Viterbi algorithm [1]–[5].

The mapping table in Figure 2 maps the output binary M-tuples code symbols from an $R = m/n$ convolutional code into binary M-tuples, which in this paper are codewords from a spectral nulls codebook. To illustrate this idea, we first present an example. We use the simple, "generic text book example" four state, binary $R = 1/2$, $v = 2$, convolutional code with octal generators 5 and 7 [6]. At the output of the encoder, we can map the set of binary 2-tuple code symbols, $\{00, 01, 10, 11\}$ onto a set of M-tuples spectral nulls code words, whether they are binary or non-binary codewords as will be explained in the following sections.

The property of distance increasing can be verified by setting up the matrices $D = [d_{ij}]$ and $E = [e_{ij}]$. Here d_{ij} and e_{ij} are respectively the Hamming distance between the code words of the convolutional code and the mapper code, the latter is here the spectral nulls code.

In general, three types of DPMs can be obtained, depending on how the Hamming distance is preserved.

- In the case where $e_{ij} \geq d_{ij} + \delta$, $\delta \in \{1, 2, \ldots\}$, $\forall i \neq j$ we call such mappings *distance-increasing mappings* (DIMs).
- In the case where $e_{ij} \geq d_{ij}$, $\forall i \neq j$ and equality achieved at least once, we have *distance-conserving mappings* (DCMs).
- In the case where $e_{ij} \geq d_{ij} + \delta$, $\delta \in \{-1, -2, \ldots\}$, $\forall i \neq j$, we have *distance-reducing mappings* (DRMs).

Example: We take $\{010101, 101010, 011100, 100011\}$ as 6-tuples, which the corresponding state systems appearing in Figure 3. In our example we have:

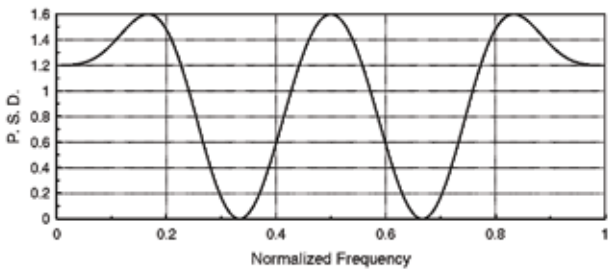


Figure 1: PSD of SN code for $M = 6$ and $k = 3$.

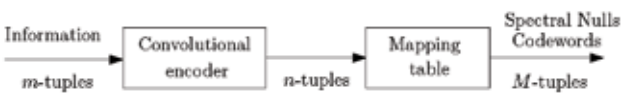


Figure 2: Encoding process for a distance-preserving spectral nulls code

$$D = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 0 & 1 \\ 2 & 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad E = \begin{bmatrix} 0 & 2 & 4 & 6 \\ 2 & 0 & 6 & 4 \\ 4 & 6 & 0 & 2 \\ 6 & 4 & 2 & 0 \end{bmatrix}. \quad (5)$$

We can see that for all i, j, $i \neq j$, $e_{ij} \geq d_{ij}$. In fact, for $i \neq j$, $e_{ij} \geq d_{ij} + 1$ and this guarantees an increase in the distance of the resulting code. The performances of the base code and different spectral nulls codes are compared in Figure 4.

As the mappings have changed the error correction performance of the resultant code compared to the base code, we can also see the change that happened on the spectrum of the code.

Figure 5 shows the spectrum of the DPM code and it is clear that the spectrum has changed compared to the original code, which is presented in Figure 1. Thus our mapping technique leads to a code with different spectrum and better error correction capability.

Finally we can say that changing the base code rate to higher rates [7], will cover more codewords from the original codebook because of the increase in the number of outputs of the convolutional codes as depicted in Table I.

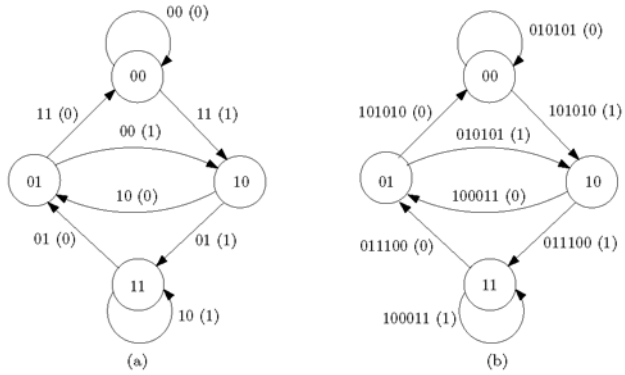


Figure 3: State systems for the (a) convolutional code and (b) DPM-SN code.

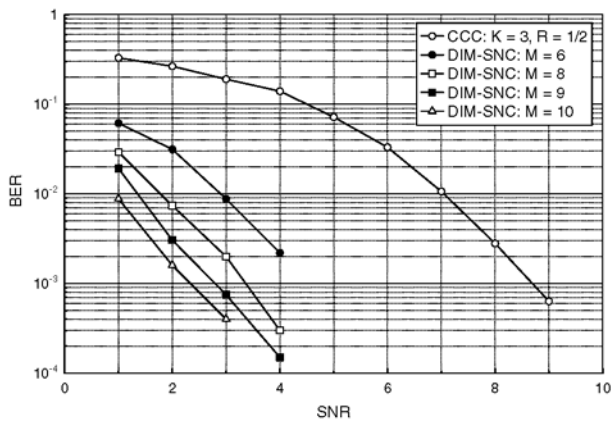


Figure 4: Hard decision decoding with QPSK/OFDM and AWGN.

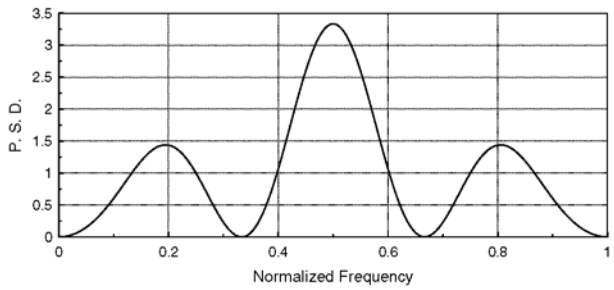


Figure 5: PSD of DPM-SN code for $M = 6$ and $k = 3$.

Table I: Variation of Output Codewords

Convolutional Code Rate	Constraint Length K	Number of Output Codewords
1/2	3	4
2/3	4	8
3/4	5	16
4/5	7	32
5/7	7	16

4. OFDM MODULATION SCHEME

In this section we do not present OFDM in detail, which can be found in the literature. We will focus on two major properties of this modulation scheme, which are related to our technique of cancellation of the narrowband interference. We will discuss the serial to parallel (S/P) stage in the OFDM scheme and the orthogonality property of this modulation.

4.1. *Serial To Parallel*

In the OFDM scheme and at the modulator we have to perform the IFFT operation to take the modulated data from the frequency domain to the time domain. The transmitted data, which is a sequence in the time domain reaches the OFDM modulator, where it will be presented in the frequency domain by the serial to parallel part in the OFDM modulation scheme as depicted in Figure 6. The loading technique of the information data into the OFDM sub-carriers bins is based on the type modulation technique that is chosen depends on the amount of data that we have to send. In our case we are using the QPSK modulation scheme [8]–[9].

4.2. *Orthogonality of the sub-carriers*

The IFFT transform is a linear transformation, which maps the complex data symbols $X(k)$, with $k \in \{0, 1, \ldots, N-1\}$, to OFDM symbols $x(n)$, with $n \in$

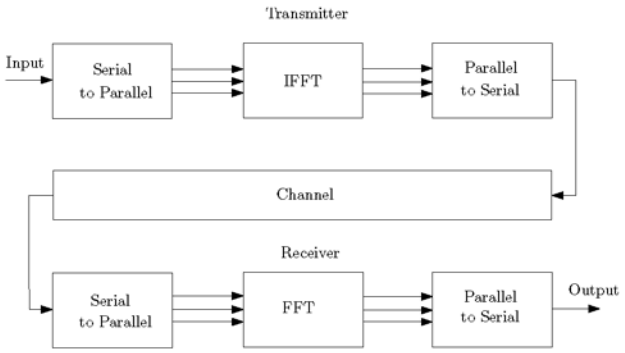


Figure 6: OFDM Block Diagram

$\{0, 1, \ldots, N-1\}$ such that,

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) e^{-j\frac{2\pi k}{N}n}. \tag{6}$$

The mapping linearity can be represented in a matrix form as:

$$x(n) = X(k) \times A_{\text{IFFT}} \tag{7}$$

where

$$A_{\text{IFFT}} = \begin{bmatrix} 1 & \ldots & 1 & 1 \\ 1 & A & \ldots & A^{N-1} \\ 1 & A^2 & \ldots & A^{2(N-1)} \\ \vdots & & & \vdots \\ 1 & A^{N-1} & & A^{N(N-1)} \end{bmatrix} \tag{8}$$

with $A = e^{-j\frac{2\pi k}{N}n}$ and $k, n \in \{0, 1, \ldots, N-1\}$.

Example: $N = 4$; $x(n) = \frac{1}{4}\sum_{k=0}^{3} X(k) e^{-j\frac{\pi}{2}kn}$, with $n \in \{0, 1, 2, 3\}$ with

$$A_{\text{IFFT}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & j & -1 & -j \\ 1 & -1 & 1 & -1 \\ 1 & -j & -1 & j \end{bmatrix} \tag{9}$$

and

$$A^H_{\text{IFFT}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -j & -1 & j \\ 1 & -1 & 1 & -1 \\ 1 & j & -1 & -j \end{bmatrix} \tag{10}$$

and this leads to the fact that $A_{\text{IFFT}} \times A^H_{\text{IFFT}} = NI$, which means that the outputs of the IFFT transform are independent and the orthogonality of the OFDM modulation is always satisfied.

5. MODIFIED OFDM MODULATOR

As presented in the previous sections, we need to take our information data from serial to parallel (S/P) to load them onto the OFDM carriers and this is what is called the serial to parallel step in an OFDM transmitter modulation scheme.

From the definition of the spectral nulls code and the algorithm of their construction, we can say that the designed codewords have specific patterns satisfying Equation 3, and their spectrum has nulls at the specific designed frequencies.

To explain this better, we suppose that a designed spectral nulls codebook S_N has a null at the frequency f_N. This means that the power spectrum density is zero at the frequency f_N. We can say that if we load these codewords from the codebook S_N onto a carrier with a frequency equal to f_N, we will have no power at that specific frequency. Thus these codewords should keep their pattern safe as it was designed to guarantee a null at the specific frequency.

From Figure 7 (a), we can say that the S/P step in the OFDM transmitter may disturb that pattern and then we will not be able to have the same sequence of data that will create a null at the specific frequency.

A modification to the OFDM modulator is necessary to overcome this problem. Figure 7 (b) shows the modified OFDM transmitter where an OFDM multiplexer is added and has the role of rearranging the data in such a way when loaded on carriers, that we can preserve the sequence and patterns of the spectral nulls data.

As example if we send the following data stream, which is generated from our spectral nulls codebook, **000111** 001110 111000 110001, the OFDM multiplexer will arrange it as shown in Table II.

As we can see, our multiplexer arranges our data column by column from top to bottom and then read data out row by row from left to right (see the bolded-font bits).

We can see from Table II that the number of columns is equal to the number of carriers and with this way of rearranging our data, when we pass via the S/P stage in the OFDM, we will guarantee that our data loaded

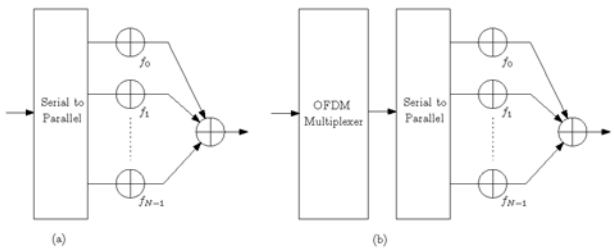


Figure 7: (a) Standard OFDM Modulator. (b) Modified OFDM Modulator.

Table II: OFDM Multiplexer

OFDM-Multiplexed Data				
	cw_1	cw_2	cw_3	cw_4
Codeword of length M	**0**	0	1	1
	0	0	1	1
	0	1	1	0
	1	1	0	0
	1	1	0	0
	1	0	0	1
	Number of Carriers			

onto the carriers preserve the sequence of the designed codebook.

6. CANCELLATION AND CORRECTION OF NARROWBAND INTERFERENCE

6.1. *Channel description for narrowband noise*

The narrowband noise signal has an energy E_{Ms} and a frequency $f_{Ms} = f_i$, where f_i corresponds to one of the transmission frequencies. This energy exceeds the received power level by orders of magnitude. We assume that the narrowband noise source has the probability ρ_n to be present, and the duration will be for $M \times t$ symbols, where $t \in [0, \infty)$. The assumption that the duration of the narrowband noise is a multiple of the duration of a codeword in our spectral null codebook is to simplify the problem. However, if asynchronous transmission is used, the duration of the narrowband noise disturbance is likely to exceed the duration of the transmission making the above assumption valid.

6.2. *OFDM in the Presence of Narrowband Noise*

Narrowband signals are considered to be approximately in the range of a single carrier [10]–[11]. The output energy of the narrowband noise source always exceeds the output energy of the modulator, which will cause the demodulator to have a consistent symbol that is *always on*, and this means that the outputs at the demodulator will be saturated to the maximum value of energy as we will prove in the following proposition.

Proposition: *The presence of a narrowband signal causes the saturation of the energy at the corresponding single carrier.*

Proof: To simplify the concept we suppose that no multipath channel exists, so we consider our channel as the combination between the AWGN (GN) channel and the Narrowband noise (NBN), so the transmitted data X_{tx} will be affected by both noises and the received data X_{rx}, will be:

$$\begin{aligned}[X_{rx}] =& [X_{tx}] + [GN] + [NBN] \\ =& [X_{(tx+GN)}] + [NBN]\end{aligned} \tag{11}$$

We can see that the term (NBN) contributes to the received modulation symbols of all subchannels as a second noise besides the additive Gaussian noise.

Taking into consideration the linearity of the FFT/IFFT transform we can apply this property to Equation 12:

$$\begin{aligned}\mathrm{FFT}(X_{rx}) &= \mathrm{FFT}([X_{tx}] + [GN] + [NBN]) \\ &= \mathrm{FFT}(X_{(tx+GN)} + NBN) \\ &= \mathrm{FFT}(X_{tx+GN}) + \mathrm{FFT}(NBN)\end{aligned} \tag{12}$$

Considering the narrowband noise as a single OFDM subcarrier, will give this noise the sinusoidal form and by looking at the spectrum of this narrowband noise after applying the Fourier transform $\mathrm{FFT}(NBN)$ as presented in Equation 12, we can say that we get an additional pulse presented by additional energy to the received signal at the corresponding subcarrier and this will cause the received data at the demodulator to be *always on*, which means we get the all "000...000" or "111...111" codewords, depending on the length of the codewords of the original spectral nulls codebook as explained previously.

6.3. *Detection and cancellation of the narrowband interferences*

Detection of the NBI:

Designed spectral nulls codebooks contain the codewords $\overbrace{00\ldots0}^{M}$ and $\overbrace{11\ldots1}^{M}$, so to be able to trace and detect the NBI, we design a coset of the spectral nulls codebook, which does not contain these two codewords. We only chose our mapping codewords from this coset to avoid having $\overbrace{00\ldots0}^{M}$ and $\overbrace{11\ldots1}^{M}$ and this make it clear and easy to detect the presence of a NBI.

At the demodulator we have to add a Power Spectrum Density (PSD) detector or a NBI detector such as in [11], which can detect the disappearance of the Nulls at the specific frequencies caused by the appearance of the NBI. Figure 8 shows the modified demodulator where a PSD detector is added.

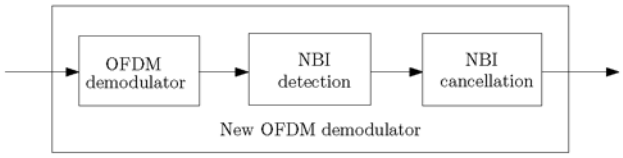


Figure 8: New OFDM Demodulator.

Cancellation of the NBI:

The cancellation step comes after the detection of the NBI as shown in Figure 8. Canceling narrowband interference by creating a null at the corresponding frequency is the same as having the sequence of our codewords with the same patten of a designed spectral nulls codewords at the specific frequency.

At the demodulator, the data will will be located in the time-frequency grid, which is similar to a matrix with the number of rows equal to the number of carriers. Once we have located the all zeros or ones cause by the narrowband interference for multiple of the length of the codeword M, we cancel it by forcing that corrupted sequences to have the designed codebook sequence, which is the mapper code. The technique of the look-up table used here to cancel narrowband interference is actually based on the search of the minimum Hamming distance between the corrupted codeword and any element of the mapper code. The corresponding codeword will replace the corrupted one.

The power spectrum density of the substituted sequence will have nulls at the specific frequencies of the designed codebook. The technique of cancellation is general and can be implemented to any type of channels where the narrowband interference is highly present [12].

To guarantee a perfect detection and cancellation of the narrowband interference, it is advisable to have the nulls as wide as possible and this can be achieved by two methods. We can can have different spectral nulls codebooks which generate the same nulls but with different length of codewords and this results of the differences in the values of the s, which refer to the number of symbols in each grouping.

Figure 9 shows how the wideness of a nulls vary with the variation of the parameter s.

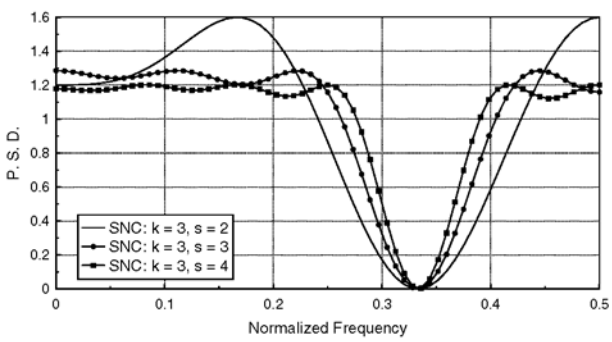


Figure 9: PSD of SN code for $k = 3$ with different values of s.

The second method is to choose different codewords for the mappings with different digital sum variations. We have chosen three sets used for the mapping and plot their spectrum. $S_1 = \{010101, 011100, 100011, 101010\}$, $S_2 = \{000111, 010101, 101010, 111000\}$, $S_3 = \{000111, 001110, 110001, 111000\}$. Figure 10 shows the differences of the wideness of three sets of different codewords from the same original designed spectral nulls codebook.

6.4. *Error correction of the Narrow band Interference*

Canceling the narrow band interference is actually not error correcting the information data. As it is known that the look-up table is in fact a limited error-correction technique, so we need to correct our received data with other error correction algorithms like the Viterbi decoder. Before that we have to demultiplex the outputs from the OFDM demodulator.

As mentioned before, the DPM technique strengthen the performance of our Viterbi decoder. In the end, we get a new convolutional code as shown in Figure 11. Figure 11 shows the block diagram of the new system including the new OFDM modulator/demodulator. The performance of the system is shown in Figure 4.

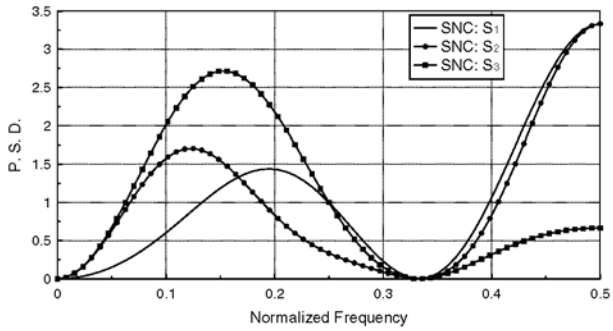


Figure 10: PSD of SN code for $M = 6$, $k = 3$ with different subsets.

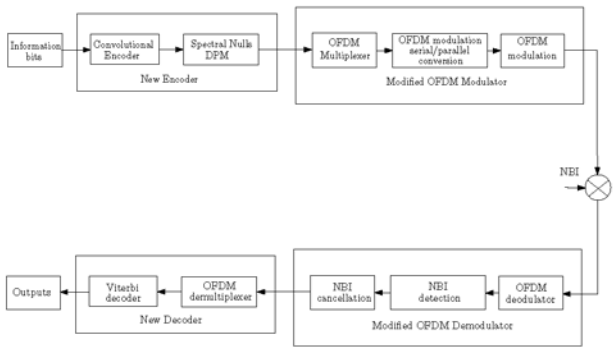


Figure 11: Block Diagram of the New Error Correction System

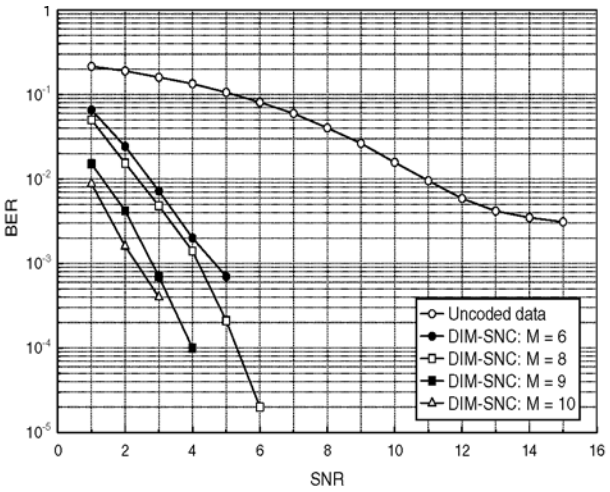


Figure 12: Hard decision decoding with QPSK/OFDM and AWGN + NBI.

It is clear from Figure 12 that our technique of DPM and the multiplexing of our data at the input of the OFDM transmitter helped us to track, cancel, and correct the data corrupted by the NBI.

7. CONCLUSION

New techniques of cancellation and correction of NBI were introduced. We have achieved three new things. The first is the generation of a spectral nulls codes, which previously used to be block codes, and which are now obtained from a convolutional code by using the DPM technique. The second is the improvement of the performance of the error correction capabilities of the Viterbi decoder by increasing the minimum distance. Thirdly a new OFDM transmitter and receiver were designed to be suitable to cancel the NBI.

The Viterbi decoder has a limitation against burst errors and NBI usually causes such errors, so we have to look further at how we can strengthen our decoder to be able to combat this type of errors better.

On the other hand we have to utilize the original spectrum as much as we can and this is done by using more codewords from the mapper code. Here we have to use convolutional codes with higher rates like 2/3 or 3/4 to map more codewords. We are also investigating another technique to better cover the spectrum and this by using the time-varying technique, which allows us to employ more codewords taking into consideration the changes that occur with the Viterbi Algorithm using a time varying trellis diagram.

ACKNOWLEDGEMENT

This material is based upon work supported by National Research Foundation under Grant Number 2053408, and furthermore partly completed during a visit at the University of Duisburg-Essen, Germany, under the supervision of Prof. A. J. H. Vinck.

8. REFERENCES

[1] H. C. Ferreira, A. J. H. Vinck, T. G. Swart, and I. de Beer, "Permutation trellis codes," *IEEE Trans. Commun.*, vol. 53, no. 11, pp. 1782–1789, Nov. 2005.

[2] E. Gorog, "Redundant alphabets with desirable frequency spectrum properties," *IBM J. Res. Develop.*, vol. 12, pp. 234–241, May 1968.

[3] K. A. S. Immink, *Codes for mass data storage systems*, Chapter 12, Shannon Foundation Publishers, The Netherlands, 1999.

[4] D. Slepian, "Permutation modulation," *Proc. IEEE*, vol. 53, no. 3, pp. 228–236, March 1965.

[5] T. G. Swart, I. de Beer, H. C. Ferreira, and A. J. H. Vinck, "Simulation results for permutation trellis codes using M-ary FSK," *Proc. Int. Symp. on Power Line Commun. and its Applications*, Vancouver, Canada, April 6–8, 2005, pp. 317–319.

[6] A. Viterbi and J. Omura, *Principles of digital communication and coding.* McGraw-Hill Kogakusha LTD, Tokyo Japan, 1979.

[7] D. G. Daut, J. W. Modestino, and L. D. Wismer, "New short constraint length convolutional code constructions for selected rational rates," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 5, pp. 794–800, September 1982.

[8] J. G. Proakis and M. Salehi, *Communication systems engineering.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1994.

[9] V. K. Ingle and J. G. Proakis, *Digital signal processing using Matlab V.4.* PWS Publishing Company, Boston, 1994.

[10] D. Galda and H. Rohling, "Narrow band interference reduction in OFDM-based power line communication systems," *Proc. Int. Symp. on Power Line Commun. and its Applications*, Malmo, Sweden, April 4–6, 2001, pp. 345–351.

[11] R. Nilsson, F. Sjoberg, and J. P. LeBlanc, "A rank-reduced LMMSE canceller for narrowband interference suppression in OFDM-based systems," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2126–2140, December 2003.

[12] K. Ouahada, H. C. Ferreira, A. J. H. Vinck, and W. A. Clarke, "Combined higher order spectral nulls codes and OFDM modulation for power line communications," *Proc. Int. Symp. on Power Line Commun. and its Applications*, Orlando, Florida, USA, March 26–29, 2006, pp. 122–127.

SOFTWARE AIDED DESIGN OF A CMOS BASED POWER AMPLIFIER DEPLOYING A PASSIVE INDUCTOR.

M. Božanić and S. Sinha

CEFIM, Dept. of Electrical, Electronic and Computer Engineering, University of Pretoria, Lynnwood Road, Pretoria, 0002, South Africa

Abstract: This paper presents the design methodology of an integrated power amplifier, and coins the methodology as a software routine: for a given set of power amplifier specifications and CMOS process parameters, the routine computes the passive component values for a Class-E based power amplifier. The routine includes the matching network for standard impedance loads. The program also provides its user with a spiral inductor calculator, which can be used to determine inductance and parasitic values for an integrated square planar spiral inductor. The same tool has the ability to extract a Simulation Program With Integrated Circuit Emphasis netlist of inductor geometry, which can be used in the post-layout simulations of the power amplifier. Operation of the program was demonstrated by simulations in Austria Microsystems 0.35 μm single-supply process for a 10 dBm, 2.4 GHz power amplifier design.

Key words: Power amplifier, spiral inductor, Class-E amplifier, impedance matching, SPICE netlist.

1. INTRODUCTION

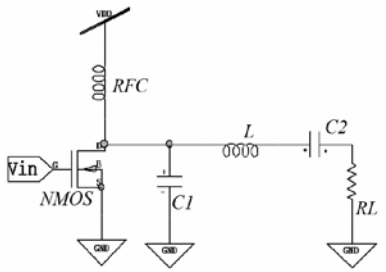
Power amplifiers (PAs) remain a bottleneck in the full integration of wireless transceivers, especially if integration is done in pure Silicon CMOS processes. For this reason, most commercial solutions for wireless LAN (WLAN) use an external PA to drive an antenna. This external device is usually fabricated in technologies superior to Silicon CMOS, such as Silicon-Germanium, Gallium-Arsenide, Indium-Phosphate or Gallium-Nitride processes. Additionally, the integrated PA disturbs analogue signals if it is fabricated on the same IC, and its performance is inferior compared to the external PA devices.

Nevertheless, it is not always possible to separate a PA from the rest of the system, so the designers often have to be satisfied with a simple PA design in a pure Silicon CMOS process.

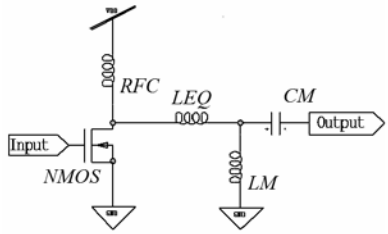
In many radio frequency (RF) modulation schemes, such as the Direct Sequence Spread Spectrum (DSSS) technique, it is necessary to design several PAs in order to ensure operation in different channels of the same band. For example, for a system based on DSSS [1], transmission is possible over a number of channels in the Industrial, Scientific and Medical (ISM) band; their centre frequencies ranging from 2.4000 GHz to 2.4835 GHz. If the design procedure is not aided by a computer program, designing of all PAs for all channels can be quite time consuming.

Irrespective of the configuration of the PA used in the design, in addition to active components (MOSFET transistors), a number of passive components (inductors and capacitors) should be included. In SPICE (Simulation Program With Integrated Circuit Emphasis) simulations at schematic design level this does not present a problem, because ideal capacitors and inductors can be used. However, translation of the design onto layout level (where layout refers to actual drawing of components for a silicon wafer) becomes more intricate. Each capacitor and inductor must be designed separately using different silicon and/or metal layers. After completing the design of these components, their netlists are extracted, and ideal components from the schematic level are then replaced with non-ideal layout-drawn components.

In the case of a capacitor, the layout tool is provided with built-in procedures for its SPICE netlist extraction, and the actual capacitance of drawn components can be checked. If necessary, this value can be fine-tuned by slightly modifying the geometry of the component. Unfortunately, in more affordable electronic design automation (EDA) software, for instance Tanner Tools, this does not apply to inductors. The built-in SPICE netlist extractor "sees" the inductor as simply a long piece of metal or wire (which short-circuits the nodes placed at the two ends of an inductor), thus making the extracted netlist incorrect. In order to perform post-layout simulations, this can be rectified by inserting "ideal" inductors between the short-circuited nodes, which in turn leads to incorrect simulation results. Such an approach becomes particularly inappropriate in the design of PAs, due to the fact that even small differences between actual and designed values of inductance can strongly affect the centre frequency or the gain of the amplifier. Often these mismatches can only be seen after the fabrication of the chip is completed.



(a) Schematic presentation of a PA [3]



(b) PA with components L_{EQ}, C_M and L_M calculated by the PA design program

Figure 1: Designed CMOS power amplifier

In this paper, a computer program is presented to simplify the design process of integrated PAs and inductors associated with them, especially in cases where time is at stake. A number of parameters, such as centre frequency, required output power and the power supply voltage are taken as inputs to this program. Values of the passive components are then calculated and optimized. The program also performs impedance matching for standard antenna impedances (50 Ω).

The program offers an option to calculate inductance and parasitic values for given spiral inductor geometries in order to aid the designer in the layout design. Finally, the netlist for such inductor geometry can be extracted. The designer can then place this netlist directly into the netlist of the global system extracted by SPICE to adequately compensate for the abovementioned short-circuit problem.

2. POWER AMPLIFIER DESIGN EQUATIONS

A number of simple PA output stages are available in modern electronics. Some of these stages, such as Class A, Class AB or Class C are commonly used in power electronics. For RF ICs, Class-E and Class-F stages are more suited [2].

The PA program presented here computes component values for the Class-E stage, as shown in Figure 1(a) [2]. This configuration is powered from a single power supply, V_{DD}.

If the required output power is labelled as P_O, the optimum load resistance can be calculated as [3]:

$$R_L = 0.577 \frac{V_{DD}^{\ 2}}{P_O} \tag{1}$$

Furthermore, if Q is defined as quality factor (Q-factor) of the PA by letting

$$Q = \omega_0 / BW \tag{2}$$

where:

BW = transmission bandwidth
ω_0 = centre transmission frequency in rad/s

then values for three passive components from Figure 1(a) (one inductor and two capacitors) can be calculated as follows:

$$C_1 \approx \frac{1}{5.447 \omega_0 R_L} \tag{3}$$

$$L = \frac{Q R_L}{\omega_0} \tag{4}$$

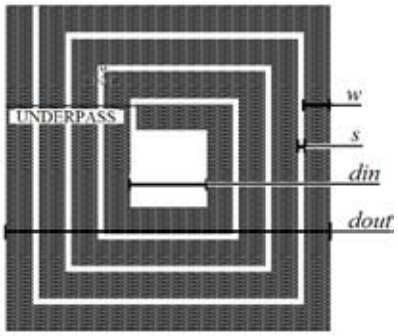
and:

$$C_2 \approx C_1 \frac{5.447}{Q}\left(1 + \frac{1.42}{Q - 2.08}\right) \tag{5}$$

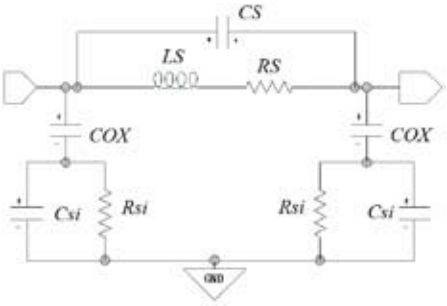
In most schematic designs, capacitor C_1 is omitted, because this value is on the same order as the value of parasitic drain-to-source capacitance of the transistor. In general, the parasitic effects of the transistor are not taken into account by the PA design program, as they are process specific. It is assumed that the designer will be responsible for further fine-tuning a PA for its correct operation.

Equations 3, 4 and 5 are sufficient for the design of a PA if it is driving the antenna with impedance calculated in Equation 1. However, this value of R_L is a non-standard load impedance, so impedance matching to the actual load has to be done. The PA program performs this matching by introducing a high pass reactive-L-type matching network, with an inductor (L_M) connected between the output of the PA and ground, and a capacitor (C_M) connected between the output of the PA and the load. This configuration is chosen specifically so that capacitor C_M can at the same time serve as decoupling capacitor for the load. Since both output impedance of the PA and the impedance of load are real values, a complex matching procedure, traditionally performed on Smith charts, can be simplified to evaluating the following two real formulas [4]:

$$L_M = \frac{1}{\omega_0}\left(\frac{R_O^{\ 2} R_M}{R_O - R_M}\right)^{1/2} \tag{6}$$



(a) Layout of the spiral inductor



(b) Model of the spiral inductor [6]

Figure 2: Layout and model of the spiral inductor

and:

$$C_M = \frac{1}{\omega_0}\left(\frac{R_O^{\,2} + (\omega_0 L_M)^2}{R_O^{\,2}(\omega_0 L_M)}\right) \tag{7}$$

Finally, due to the very high value of the calculated Q-factor, the value of the calculated inductor L is higher than the practical value, whilst the value of the capacitor C_2 is lower than the practical value. These two components appear in series, and more practical values can be obtained applying the procedure described in [5]. If Z_L is the impedance of the inductor L and Z_{C2} the impedance of capacitor C_2, then the equivalent impedance is:

$$Z_{EQ} = Z_L + Z_{C2} \tag{8}$$

This results in a single inductor of value (L_{EQ}) practical for microelectronic integration.

Figure 1(b) shows the final PA circuit, with all components calculated using the PA design program. This program does not suggest the transistor aspect ratio (W/L). However, to ensure that the output stage has voltage gain larger than 1, this ratio should be kept as high as possible. For the Austria Microsystems (AMS) CMOS 0.35 µm process, this value was chosen to be the highest allowed, (200 µm)/(0.35 µm).

3. SPIRAL INDUCTOR DESIGN EQUATIONS

Demand for low-cost ICs has generated high interest in integrated passive components. Whilst capacitor and resistor implementations are easy to model, as discussed before, considerable effort has been invested in the design of integrated inductors [6]. Two practical options are bond wires and planar spiral inductors. Bond wires allow for high Q-factor of the inductor to be achieved, where Q-factor is defined as Q = Im(Z)/Re(Z) and Z is the impedance of the inductor, which is different from the PA Q-factor described in Section 2. However, the range of possible inductances for bond wire inductors is limited, leaving spiral inductors as the only other feasible option for integration. Although the Q-factor of the spiral inductor is lower than the Q-factor of the bond wire inductor, its inductance is well-defined over a broader range of process variations.

A number of spiral inductor configurations are commonly used, amongst them square, hexagonal, octagonal and circular. Figure 2(a) shows the layout of a square inductor. Because of the ease of its layout, the PA design program suggests this configuration for implementing inductors for the designed PA.

Although the idea behind the process of designing a spiral inductor is to get as close as possible to the desired inductance values, some parasitic effects cannot be avoided. For the purpose of expanding the PA design program, the inductance value, and not the Q-factor, is considered important for the design. Parasitic values are still calculated, and included in the netlist. Figure 2(b) shows the circuit model for the spiral inductor.

A spiral inductor is fully specified by the number of turns (n), the turn width (w) and two of the following: inner, outer or average diameter (d_{in}, d_{out} or $d_{avg} = (d_{in}+d_{out})/2$), as shown in Figure 2(a). The spacing between the turns (s) can be calculated from the other parameters.

3.1 Inductance (L_S)

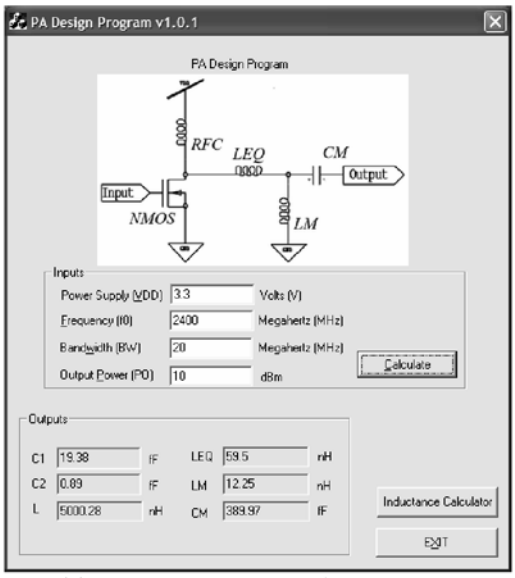
Inductance of the square spiral inductor is calculated by means of a data fitted monomial expression [6][7] (with inductor dimensions in micrometer):

$$L_S = \beta d_{out}^{\alpha_1} w^{\alpha_2} d_{avg}^{\alpha_3} n^{\alpha_4} s^{\alpha_5} \text{ (nH)} \tag{9}$$

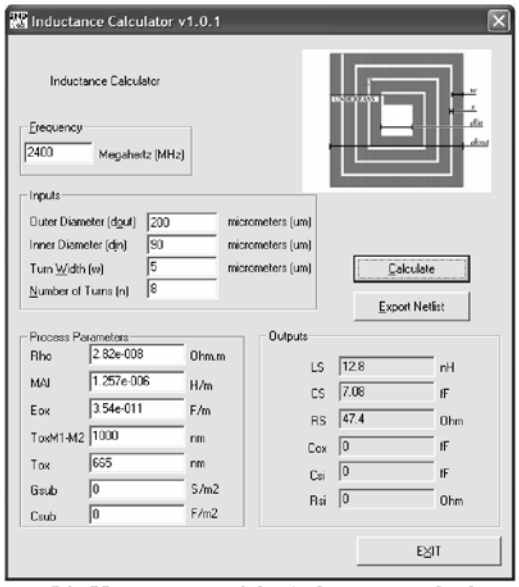
Above, the coefficients β, α_1, α_2, α_3, α_4 and α_5 are obtained numerically and presented in Table I [6]. This expression yields the typical error of less than 3%.

Table I: Coefficients for square spiral inductor inductance calculation

β	α_1 (d_{out})	α_2 (w)	α_3 (d_{avg})	α_4 (n)	α_5 (s)
$1.62 \cdot 10^{-3}$	-1.21	-0.147	2.40	1.78	-0.030



(a) User screen of the PA design program



(b) User screen of the inductance calculator

Figure 3: User screens and of PA design program and inductance calculator.

3.2 Parasitic resistance (R_S)

The parasitic resistance is dependent on the frequency of operation. At DC, this value is mostly dependent on sheet resistance of the material of which the wire is made. At high frequencies, this is overshadowed by the resistance that arises due to formation of eddy currents. It depends on the resistivity of the metal layer in which the inductor is laid out (ρ), total length of all inductor segments (l), width of the inductor (w) and its effective thickness (t_{eff}) [8]:

$$R_S = \frac{\rho l}{w t_{eff}} \tag{10}$$

Effective thickness, t_{eff}, is dependent on the actual thickness of the metal layer, t:

$$t_{eff} = \delta(1 - e^{-t/\delta}) \tag{11}$$

where δ is frequency-dependent skin depth:

$$\delta = \sqrt{\frac{\rho}{\pi \mu f}} \tag{12}$$

Parameter μ is the permeability of the metal layer.

3.3 Parasitic capacitance (C_S)

The parasitic capacitance is the sum of all overlap capacitances created between the spiral and the underpass. If the underpass has the same width as the spiral, then the capacitance is equal to [8]:

$$C_S = nw^2 \frac{\varepsilon_{ox}}{t_{oxM1-M2}} \tag{13}$$

where:
$t_{oxM1\text{-}M2}$ = oxide thickness between the spiral and the underpass
ε_{ox} = dielectric constant of the oxide layer between the two metals.

3.4 Substrate parasitic quantities (C_{ox}, C_{Si} and R_{Si})

Substrate parasitic quantities are approximately proportional to the area occupied by the inductor and can be estimated as [8]:

$$C_{ox} = \frac{1}{2} lw \frac{\varepsilon_{ox}}{t_{ox}} \tag{14}$$

$$C_{Si} = \frac{1}{2} lw C_{sub} \tag{15}$$

and:

$$R_{Si} = \frac{2}{lwG_{sub}} \tag{16}$$

where:

C_{sub} = capacitance per unit area for the silicon substrate
G_{sub} = conductance per unit area for the silicon substrate
ε_{ox} = dielectric constant of the oxide layer between the inductor and the substrate
t_{ox} = thickness of the oxide layer.

These expressions are sufficient for modelling the spiral inductors for the requirements of the PA design program [7].

4. THE PROGRAM

The PA design program was written in Visual C++.

4.1 Power amplifier design part of the program

The primary screen of the PA design program is shown in Figure 3(a). This screen is used to calculate values of the PA passive components. The user inputs the power supply voltage (V_{DD}) in volts, required output power (P_O) in dBm, centre frequency of the channel over which the amplifier operates (f_0) in MHz and allowed bandwidth (BW) in MHz. Values of capacitors C_1, C_2 and inductors L and L_{EQ}, together with matching parameters C_M and L_M are then computed and displayed. Although L_{EQ}, C_M and L_M are sufficient for the complete simulation using ideal passive components, the non-simplified values C_1, C_2 and L are available for experimentation by the user.

4.2 Spiral inductor design part of the program

This part performs the secondary function of the program. In order to design real spiral inductors, the user is allowed to specify the geometry of the square inductor and see the resulting inductance. Figure 3(b) shows the user screen of the inductance calculator. The user needs to input the geometry of the inductor, namely the outer diameter (d_{out}), inner diameter (d_{in}), width of one turn (w), all in μm, and the number of turns (n). Process-specific values for calculating parasitic quantities need not be specified, as default parameters are used (for the AMS 0.35 μm process). If necessary, they can be replaced with specific technology parameters. Specifically, the program uses resistivity and permeability values for aluminium wires, as shown in Table II, as well as frequency specified by the user in the primary screen, as defaults for calculating the parasitic resistance. For the calculation of other parasitic quantities, common process values are used as shown in Table II. Finally, the user is allowed to specify G_{sub} and C_{sub} for calculations of C_{si} and R_{si}. If these parameters are omitted, calculations of parasitic values C_{ox}, C_{si} and R_{si} are not performed.

Table II: Parameters for spiral inductor parasitics calculation

ρ_{Al} [Ω·m]	μ_{Al} [H/m]	ε_{ox} [F/m]	$t_{oxM1\text{-}M2}$ [nm]	t_{ox} [nm]
$2.82\cdot10^{-8}$	$1\cdot4\pi\cdot10^{-7}$	$4\cdot8.85\cdot10^{-12}$	1000	665

Optimization of the spiral inductor is performed by modifying the dimensions or trace width of the inductor. In general, the inductance can be increased by decreasing the trace width and increasing the number of turns, which, in turn, increases the parasitic resistance R_S and subsequently decreases the Q-factor of the inductor.

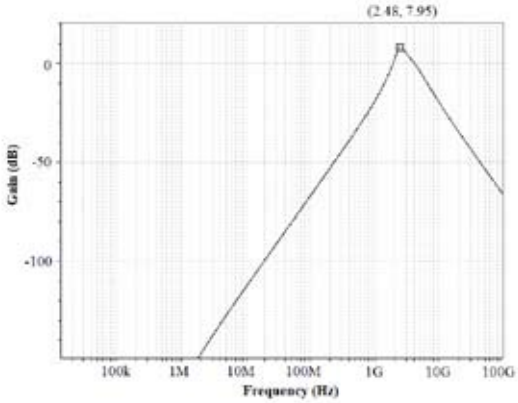
4.3 Spiral inductor netlist extraction

After a few geometry iterations, a correct inductor value can be generated, and a SPICE netlist exported (Figure 3(b)). An example of such a netlist is shown in Figure 4.

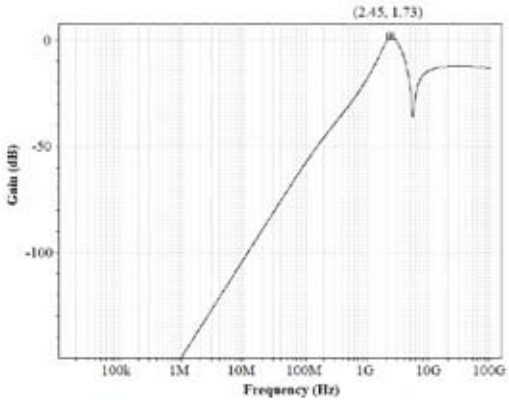
```
*Include the following code before Main Circuit
*SUBCKT gen. by Ind. Calc. v1.0.0
.SUBCKT Inductor_Model L1 L2 GND
CS L1 L2 1fF
Csi1 N2 GND 1fF
Csi2 N3 GND 1fF
Cox1 L1 N2 1fF
Cox2 L2 N3 1fF
LS N4 L1 1n
Rsi1 N2 GND 10
Rsi2 N3 GND 10
RS N4 L2 10
.ENDS

*Include the following line in Main Circuit
*instead of your ideal inductor
*and replacing nodes L1 and L2 with original
ones
XInductor_Model_1 L1 L2 IndModLEQ
```

Figure 4: Example of SPICE netlist of spiral inductor model.



(a) Bode plot of the gain of PA employing ideal inductors



(b) Bode plot of the gain of PA employing a spiral inductor designed per Figure 3

Figure 5: Simulation Bode plots

5. SIMULATION RESULTS

SPICE simulations of the extracted netlists were performed in Silicon CMOS 0.35 μm process from AMS. This process was used to verify operation of a PA, employing both ideal inductors (infinite Q-factor) and spiral inductors (finite Q-factor). This process is powered from a typical single power supply of 3.3 V.

```
...
C1 N3 N2 390fF
L2 VDD N5 50n *RFC
L3 N5 N4 59.50n *LEQ
L4 N3 GND 12.25n *LM
XPowerSupply_1 GND MST_RST VDD PowerSupply
R5 N2 GND 50
XRFNMOS7 N5 N9 GND GND MODNRF W=200u L=0.35u
NG=40
v8 N8 GND sin(0.9 1 2.4E9 0 0 0)
v9 N9 N8 AC 0.2 0
...
```

(a) Snippet of the netlist of a PA with ideal inductors

```
...
*Include the following code before Main Circuit
*SUBCKT gen. by Ind. Calc. v1.0.0
.SUBCKT IndModLM L1 L2
CS L1 L2 7.08fF
LS N1 L1 12.8n
RS N1 L2 47.40
*Ignoring Csub,Gsub
.ENDS

*Include the following code before Main Circuit
*SUBCKT gen. by Ind. Calc. v1.0.0
.SUBCKT IndModLEQ L1 L2
CS L1 L2 14.16fF
LS N1 L1 59.09n
RS N1 L2 130.8
*Ignoring Csub,Gsub
.ENDS

...
C1 N3 N2 390fF
*Include the following line in Main Circuit
*instead of your ideal inductor
*and replacing nodes L1 and L2 with original
ones
XIndModlLEQ_1 N5 N4 IndModLEQ
*Include the following line in Main Circuit
*instead of your ideal inductor
*and replacing nodes L1 and L2 with original
ones
XIndModLM_1 GND N3 IndModLM
L2 VDD N5 50n *RFC
XPowerSupply_1 GND MST_RST VDD PowerSupply
R5 N2 GND 50
XRFNMOS4 N5 N9 GND GND MODNRF W=200u L=0.35u
NG=40
V8 N8 GND sin(0.9 1 2.4E9 0 0 0)
V9 N9 N8 AC 0.2 0
...
```

(b) Snippet of the netlist of a PA with ideal inductors replaced with inductor netlists generated by inductance calculator part of the program

Figure 6: Netlist snippets

In the simulation example, it was assumed that power amplification was done for an analogue signal operating in ISM band, with centre frequency of 2.4 GHz and channel width of 20 MHz. Although for the schemes such as DSSS, centre frequencies are just above 2.4 GHz and channel widths are 22 MHz, these parameters are sufficient for the simulations of interest. Further, the required output power of the stage was assumed at 10 dBm. Figure 5(a) shows the Bode plot of the PA gain, using a 50 ? antenna for a SPICE simulation performed on a netlist (with "ideal" inductors). The snippet of this netlist is shown in Figure 6(a). Figure 5(b) shows the same plot, but with ideal inductors in the netlist of Figure 6(a) replaced by spiral inductors, designed by the secondary function of the PA design program. This presents an emulation of the post-layout simulation. The snippet of this netlist is shown in Figure 6(b). Table III presents the summary of technical parameters, PA component values, inductor geometry parameters, as well as measured parameters for the simulated circuits.

Table III: Inputs, outputs and parameters used in PA simulations

Technical Parameters		Design Parameters (Spiral Inductor)		
Parameter	Value	Parameter	Value (L_{EQ})	Value (L_M)
f_0 [GHz]	2.4	d_{out} [μm]	300	200
BW [MHz]	20	d_{in} [μm]	100	90
P_{out} [dBm]	10	w [μm]	5	5
V_{DD} [V]	3.3	n	16	8
Calculated Component Values (PA)		Calculated Inductance and Parasitics (Spiral Inductor)		
Parameter	Value	Parameter	Value (L_{EQ})	Value (L_M)
C_1 [fF]	19.38	L_S [nH]	59.09	12.8
C_2 [fF]	0.89	R_S [?]	130.81	47.40
L [μH]	5	C_S [fF]	14.16	7.08
L_{EQ} [nH]	59.5	C_{ox}[fF]	Ignored	Ignored
L_M [nH]	12.25	C_{si}, R_{si}	Ignored	Ignored
C_M [fF]	389.97			
Measured Parameters (Q_i ? ?)		Measured Parameters (Q_i < ?)		
Parameter	Value	Parameter	Value	
f_0 [GHz]	2.48	f_0[GHz]	2.45	
Gain [dB]	7.95	Gain [dB]	1.73	
P_{out} [dBm]	~ 9	P_{out} [dBm]	~ 4	

6. DISCUSSION

As it is evident from Table III, the centre frequency of the waveform in Figure 5(a) is higher than anticipated by calculations (Figure 3(b)). This effect is due to the parasitic capacitance of the transistor, and the lack of physical capacitor, C_1. The software routine developed in this paper aims to produce a first approximation for the design only, therefore this higher value of the centre frequency was considered acceptable. A further fine tuning of the values of other passive components, namely L_{EQ}, L_M and C_M can assist to obtain a more exact centre frequency. The output power (P_O) is close to the value designed (Figure 3(a)). A similar response is obtained after introducing the spiral inductors, as seen in Figure 5(b). However, a lower gain and lower output power are observed at the peak frequency, which can be attributed to the low inductor Q-factor. The zero at approximately 5 GHz is due to the capacitive effect of C_S of the spiral inductor L_{EQ}, whilst other parasitic elements do not have noticeable influence on the shape of the gain plot at frequencies of interest. It should also be noted that this Bode plot is accurate only for frequencies close to the design frequency, i.e. 2.4 GHz, because of the frequency dependency on the parasitic resistance, R_S. In general, it is evident that simulation results with parameters calculated by the PA design program shows good agreement with the PA's theoretical behaviour.

7. CONCLUSION

In this paper, a PA design program for ICs has been presented. The program can be used for three main purposes. Firstly, to determine parameters for Class-E PA output stages based on the set of equations presented in [2] and [3]. It also performs otherwise tedious impedance matching to standard impedance loads, such as 50 Ω, whilst providing suggestions on how to simplify some component values.

Secondly, the program can be used to calculate inductance and parasitic values for the square planar spiral inductor based on equations from [6] and [8]. Although the user is allowed to modify geometry to improve the quality factor of such an inductor, the main concern of this part of the program is to allow the user to get as close as possible to the required inductance value.
Thirdly, the program can be used to export a SPICE netlist of spiral inductor geometry to be used in post-layout simulations. This is considered essential, since most SPICE simulators interpret an integrated spiral inductor as a long stretch of wire (or a resistor).

To verify the parameters computed by the program, simulations, using the 0.35 µm process, were performed for designing a 10 dBm PA operating in ISM band with centre frequency of about 2.4 GHz. Two separate configurations were simulated: with ideal inductors and by using the spiral inductor program of this paper. The two simulations compared well.

8. REFERENCES

[1] N. Naudé, M. Božani and S. Sinha: "Analogue CMOS Direct Sequence Spread Spectrum Transceiver with Carrier Recovery Employing Complex Spreading Sequences," *Proc. IEEE MELECON*, pp. 1227-1231, 2006.

[2] C. Coleman: *An Introduction to Radio Frequency Engineering*, Cambridge University Press, USA, first edition, pp. 178-179, 2004.

[3] N.O. Sokal and A.D. Sokal: "Class E – A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers," *IEEE J. Solid-State Circuits,* Vol. SC-10, No. 3, pp. 168-176, June 1975.

[4] R. Ludwig and P. Bretchko: *RF Circuit Design: Theory and Applications*, Prentice Hall, USA, chapter 8, pp. 404-408.

[5] K.W. Ho and H.C. Luong: "A 1-V CMOS power amplifier for Bluetooth applications," *IEEE Trans. Circuits and Systems – II: Analog and Digital Signal Processing*, Vol. 50, No. 8, pp. 445-449, Aug. 2003.

[6] S.S. Mohan, M. del Mar Hershenson, S.P. Boyd and T.H. Lee, "Simple Accurate Expressions for Planar Spiral Inductances," *IEEE J. Solid-State Circuits*, Vol. 34, No. 10, pp. 1419 – 1424, Oct. 1999.

[7] S. Musunuri, P. L. Chapman, J. Zou and C. Liu, "Design issues for monolithic DC-DC converters," *IEEE Trans. Power Electronics*, Vol. 20, No. 3, pp. 639- 649, May 2005.

[8] C.P. Yue and S.S. Wong, "Physical Modeling of Spiral Inductors on Silicon," *IEEE Trans. Electron Devices*, Vol. 47, No. 3, pp. 560-568, Mar. 2000.

CONSIDERATIONS IN DESIGN OF LOW ALTITUDE LONG ENDURANCE SOLAR POWERED UNMANNED AERIAL VEHICLES.

J. Meyer*, J.A.F du Plessis*, W. Clark*, P. Ellis**

** Department Electrical and Electronic Engineering Science, University Johannesburg, P O Box 524, Auckland Park, 2006, South Africa*
*** Department Electrical and Electronic Engineering Technology, University Johannesburg, P O Box 524, Auckland Park, 2006, South Africa*

Abstract: This paper presents design considerations for a low altitude long endurance solar powered Unmanned Aerial Vehicle. The considerations addressed include the determination of the available solar power, the design of the Unmanned Aerial Vehicle wing for long endurance flights and the determination of the electrical power and energy balance of the Unmanned Aerial Vehicle. Using these considerations, a solar powered Unmanned Aerial Vehicle was designed and a long endurance flight simulated. Based on the simulation results it is in principle possible to design and build a solar powered low altitude long endurance Unmanned Aerial Vehicle.

Key words: Autonomous Flight, Low Altitude Long Endurance Flight, Solar Power, Unmanned Aerial Vehicles.

1. INTRODUCTION

The ability to fly for extended periods of time has long been a challenge for the aeronautical fraternity. These aircraft would provide an ideal platform for a host of potential users.

Applications for such an aircraft would mainly be focused around surveying and surveillance applications. Surveying applications include aerial photography and remote sensing. Surveillance applications entail civilian or military reconnaissance, border patrols and services utility patrols such as the monitoring of pipelines or power lines. Many of these applications however fall into the category of dull, dirty or dangerous (DDD) aerial work with its associated risks and cost. Unmanned Aerial Vehicles (UAVs) are potentially better suited for performing these DDD aerial activities.

Due to the low altitude long endurance (LALE) capability of these UAVs, the airframe and propulsion system are major design factors. Solar powered electrical drive systems coupled with an efficient energy storage system offers a very attractive propulsion method. The main advantage of utilizing a solar powered electrical propulsion system, above the conventional combustion type, is its potential to sustain flight without the need to carry excessive fuel loads. Electric propulsion systems also do not suffer from unwanted exhaust emissions which is detrimental in atmospheric sensing applications.

An aircraft capable of sustained flight, propelled by a zero emission renewable energy source, offers a very attractive aerial platform.

Much research is conducted in the field of solar powered electric cars, culminating in competitions such as the 2007 World Solar Challenge to be held in Australia [1]. Electrically powered aircraft is not a new concept and has been around a number of years. Solar powered flight for high altitude long endurance (HALE) aircraft such as HELIOS has recently been demonstrated [2]. For the smaller scale HALE UAVs, design aspects were reported by Colozza [3, 4]. The design of sustained solar powered electric flight on small scale LALE UAVs are however few and far between [5, 6]. The advances in composite airframe construction, improved energy storage mediums, high efficiency electrical motors and propellers as well as the availability of lightweight photovoltaic cells for power generation greatly increases the feasibility of constructing a LALE small scale UAV. The high level design requirement was to: "Design and build a small scale solar powered UAV capable of demonstrating sustained flight." In order to demonstrate sustained solar powered flight, the UAV has to fly autonomously and non-stop along a pre-defined flight path from Johannesburg to Cape Town, a distance of approximately 1400 km.

The following sections present the design considerations for a small scale solar powered LALE UAV. The intent is to provide insight into the critical design aspects of small scale solar powered UAVs for LALE applications.

2. SOLAR POWER

All the energy required for sustained flight with a solar powered electric aircraft must be acquired by the

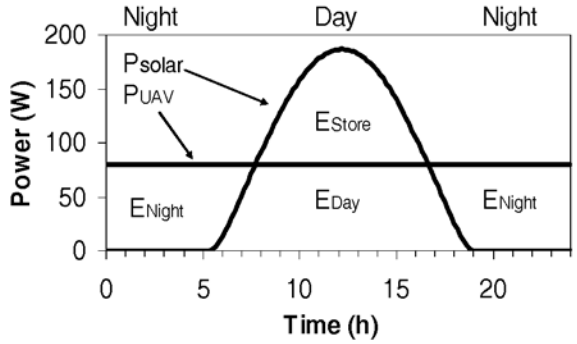


Figure 1: Energy balance diagram over a 24 hour period

photovoltaic cells fitted to the aircraft. The amount of energy to be collected during day light hours must be sufficient to propel the aircraft during the day and during the night when no additional energy is available for collection by the photovoltaic cells.

The amount of solar power available to the aircraft as a function of the time of day can be determined by considering the energy balance diagram shown in Figure 1.

Figure 1 depicts typical solar power (P_{solar}) produced by the photovoltaic cells and the power required by the UAV (P_{UAV}) to sustain flight during a 24 hour period. The energy required by the UAV for night time operation is shown as the area under the P_{UAV} curve, denoted by E_{Night}. The area shown as E_{Day}, under the P_{UAV} curve represents the energy required by the UAV during day time operation. The area denoted by E_{Store} under the P_{solar} curve represents the excess energy collected during the day light hours which is available for storage in the energy storage devices. For sustained UAV flight, the amount of energy stored during daylight hours must be greater or equal to the amount of energy required by the UAV for night time operation. The energy balance equation from Figure 1 is therefore given by

$$E_{Store} \geq 2\,E_{Night} \tag{1}$$

The total amount of energy produced by the photovoltaic cells (E_{Solar}) during the daylight hours must therefore be greater than the energy required by the UAV for day time operation and night time operation combined, as shown by the following equation:

$$E_{Solar} \geq 2\,E_{Night} + E_{Day} \tag{2}$$

The total amount of energy produced by the photovoltaic cells is a function of the UAVs geographic position (latitude, longitude, altitude), time of the year, atmospheric absorption and efficiency of the photovoltaic cells. Based on the work by Bird [7] and Kasten [8] a simple direct solar radiation model was developed. The Linke turbidity factor is used in the model to characterize the clearness of the sky. The lower this factor, the clearer the sky, the larger the beam irradiation and the lower the relative fraction of the diffuse irradiation. For higher

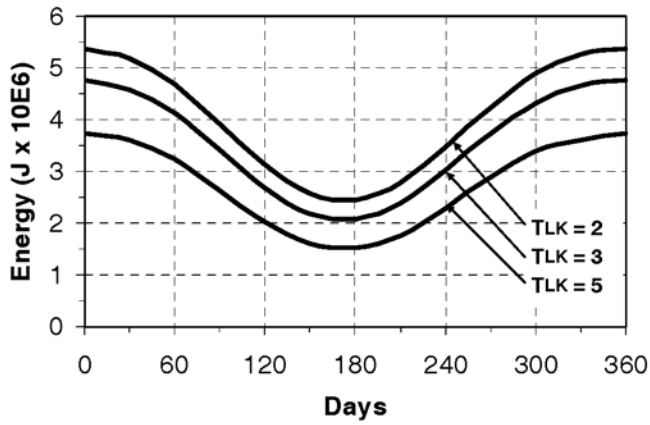


Figure 2: Available solar energy per day which can be collected by a 1m^2 photovoltaic array with a 16% efficiency for various values of Linke turbidity factor where day 1 corresponds to the summer solstice in the southern hemisphere

altitudes, the absorption is lower because of less radiation scattering by the atmosphere which lowers the Linke turbidity factor. Typical values for the Linke turbidity factor are listed in Table I.

Table I: Typical values for the Linke turbidity factor

T_{LK}	Sky Condition
1	Pure sky
2	Very clear sky
3	Clear sky
5	Summer with water vapour
7	Polluted urban industrial

Figure 2 shows the theoretically available energy which can be collected in the southern hemisphere at a latitude of 25 degrees by a photovoltaic array of 1 m^2 with an efficiency of 16% as a function of the time of the year. It follows that the amount of solar energy available per day for use in propulsion of a solar powered UAV is highly dependant on the time of the year as well as the clearness of the atmosphere. UAVs operating during the summer months, when the available energy is at its highest, have approximately 2.2 times the energy available relative to operation during winter months when the available energy is at its lowest. Even when a worst case summer day (with Linke turbidity equal to 5) is compared with a clear winter's day (with Linke turbidity equal to 2), the ratio of available solar energy in summer is still on the order of 1.5 to the available solar energy in winter. This difference in available solar energy is mainly contributed by the distance between the earth and the sun increasing in winter and the smaller sun angle and the shortening in day light hours as a result of the inclination of the earth axis.

When designing solar powered UAVs, consideration has to be given to the expected operating time of the year. Designing for minimum available solar energy conditions

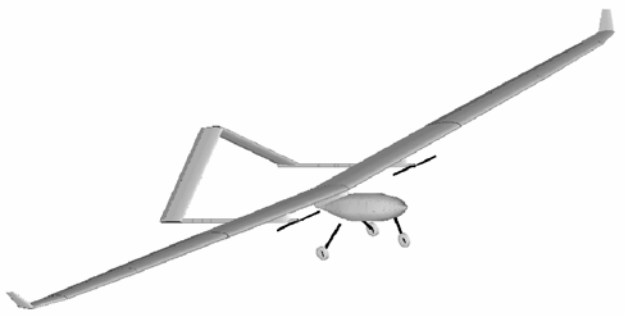


Figure 3: Twin boom inverted V tail airframe

may result in an over design by a factor of 2 under maximum available solar energy conditions. An over design factor of 2 has significant negative impact on the UAV airframe design in terms of size and cost. A positive impact may be that the excess solar energy available may be used to overcome the increased aerodynamic drag when the UAV is flying at faster speeds. This may result in UAVs which can be operated at higher flying speeds during the summer months, when the available solar energy is at a maximum.

3. AIRFRAME DESIGN

The airframe considered for the LALE solar powered UAV is illustrated in Figure 3.

3.1 Airframe Architecture

A glider type airframe was selected because of its low aerodynamic drag which is essential for low power long endurance flights. The fuselage consists of a low drag aerodynamic body mounted on a centre wing pylon. The fuselage is to house the onboard avionic systems. In order to enhance the reliability of the UAV, it was decided to use a twin motor configuration. The twin tail boom not only facilitated the mounting of the inverted V tail, but also provided for the mounting of the twin electrical propulsion motors. The inverted V tail was selected because it offers low aerodynamic drag and it also helps to reduce the loss of altitude during turn manoeuvres. Both tail booms are affixed to wing pylons. The use of under-wing-pylons to connect to the wing is a method of minimizing the perturbations of the airflow over the wing and thus minimizing the aerodynamic drag of the wing. A fixed undercarriage configuration was chosen because taking off and landing of UAVs without undercarriage requires complex and expensive launch and recovery mechanisms. Due to the undercarriage there is, however, a penalty to pay in the form of increased aerodynamic drag resulting in more power required for flight.

3.2 Airfoil Design

For the purpose of long endurance flight, a slow speed high lift airfoil is required in order to reduce the form aerodynamic drag. The total aerodynamic drag acting on a wing is the combination of the vortex drag and the form

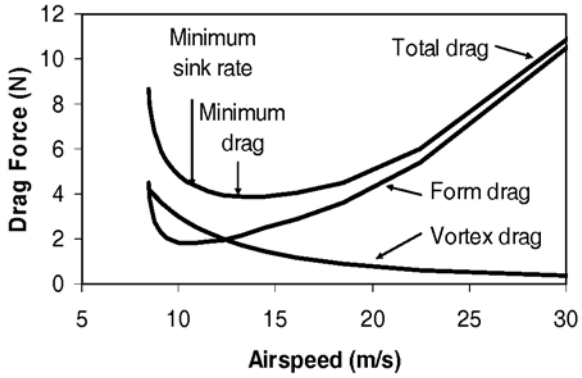


Figure 4: Aerodynamic drag forces acting on a wing section consisting of a SD7062 airfoil and wing with aspect ratio (A = 16). The drag curves are computed at Reynolds Number = 200 000, wing surface area of 1.5 m^2 and aircraft mass of 12 kg.

drag. Vortex drag is aerodynamic drag induced by the vortices trailing the wing, especially at the wing tips. Vortex drag is reduced as the speed of the aircraft increases. Form drag is caused by the total pressure variations over the wing as the air flows over it and the skin friction or viscous drag which is caused by the contact of the air with the wing surface. Form drag increases with an increase in aircraft speed. The ratio of height loss to distance covered, is the same as the glide ratio which can be shown to be equal to the lift over drag (L/D) ratio [9]. The best glide ratio is achieved at the point where the form drag is equal to the vortex drag giving the optimal L/D ratio. However, for maximum endurance the aircraft has to be operated at the minimum sinking speed. The minimum sinking speed occurs when the ratio between the coefficient of lift (C_L) and the coefficient of drag (C_D), $C_L^{1.5}/C_D$ is at its highest. This ratio is termed the power factor for the aircraft since it also indicates the trim condition for level flight with minimum thrust.

From Figure 4 it can be seen that the best lift to drag coefficient ratio (C_L/C_D) is achieved at an airspeed of

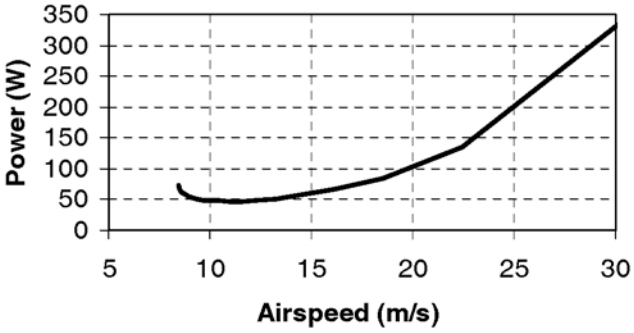


Figure 5: Power required for overcoming the drag force acting on the wing as a function of airspeed.

13 m/s. The best speed for endurance is achieved when the $C_L^{1.5}/C_D$ ratio is at a maximum, which occurs at an airspeed of 11 m/s. For long endurance flights the UAV has to operate under minimum power consumption, which is at the airspeed where the power factor is at a maximum.

The power required to overcome the total aerodynamic drag force of the wing is shown in Figure 5. It can be seen that at an airspeed of 11 m/s the minimum power required is approximately 47 W to maintain the wing in level flight. To obtain the total propulsion power required to maintain the UAV in level flight, the drag forces acting on the rest of the aircraft such as the tail and fuselage, have to be taken into account.

3.3 Wing Design

From the previous section it follows that for level flight the minimum propulsion power required is achieved when the power factor ($C_L^{1.5}/C_D$) is at a maximum. In order to maximize the power factor, the C_D coefficient must be reduced. At low airspeeds the biggest component of the drag coefficient is the vortex induced drag coefficient (C_{di}). The C_{di} coefficient is a function of the C_L coefficient, the aspect ratio (A) and the correcting factor (k = 1) of the wing as shown in Equation 3.

$$C_{di} = k(C_L^2 / \pi A) \tag{3}$$

In order to minimize the vortex drag the aspect ratio of the wing should be as large as possible. The aspect ratio of the wing is obtained by dividing the wingspan (B) by the mean cord (M),

$$A = B / M \tag{4}$$

In order to achieve the design requirement of a wing surface area of 1.5 m^2, the wing dimensions were selected as 5 m span and 0.3 m mean cord resulting in an aspect ratio of 16. Elliptical wing design with a straight trailing edge was employed to further reduce the vortex induced drag of the wing. Winglets are added to the wing tips in an attempt to reduce the wing tip vortices. To prevent wing tip stall when the UAV banks the wing tips are twisted downward to reduce its angle of attack. To enhance aircraft stability, a wing dihedral angle of 2.9 degrees was selected. The standard control surfaces consisting of ailerons and flaps are fitted to the wing. An additional benefit of adding flaps to the UAV, is that it gives the ability to change the lift versus speed characteristics of the aircraft enabling fine trim for optimal endurance.

3.4 Fuselage and Tail Design

The fuselage is typical of a low drag glider body. In order to reduce the effect of the fuselage on the airflow across the wing the fuselage is mounted on an under-wing centre pylon. The fuselage acts as housing for the avionics and some of the electrical energy storage devices. A tri-axis configuration undercarriage is fitted to the underside of the fuselage. Two tail booms are mounted under wing pylons to which the inverted tail is attached. The tail wing surface area is 0.28 m^2. A symmetrical SD8020 airfoil was selected for the tail surface.

3.5 Airframe Construction

By positioning the electrical motors and the energy storage devices on its proper positions on the wing load diagram the shear forces and bending moments of the wing were computed. Based on the maximum shear force and bending moment with a sufficient safety factor the composite material lay-up for the wing spar was computed. Wing construction is based upon a polystyrene foam core covered with composite material lay-ups. The wing spar is the main stress member and is designed with an upper and lower spar cap with a shear web on each side constructed using composite material techniques.

3.6 Design Verification

With the use of the Athena Vortex Lattice (AVL) software package [10] the geometry and the Trefftz Plane of the UAV was verified. Additionally the UAV was modelled in the Plane-Maker 8.50 [11] software package. Once the UAV was defined in Plane-maker its aerodynamic and flight behaviour was verified in the X-Plane 8.50 [11] flight simulation software package.

4. POWER MANAGEMENT SYSTEM

The propulsion and power management system is shown in Figure 6 and consists of an array of photovoltaic cells mounted on each wing, two electric motors, six lithium polymer batteries used as energy storage devices, battery charge management circuits and a power management computer.

4.1 Solar Array

The only source of electrical power for the UAV is the photovoltaic cells which are connected in an array structure. Poly-crystalline photovoltaic cells with an efficiency of 16% were selected for mounting directly on the wing surface. The cells are connected in series in order to produce the desired nominal voltage of 22.2 V.

4.2 Battery Management Systems

Each lithium polymer battery is managed by a battery management system. The purpose of the battery management system is to produce the correct voltage and current characteristics for optimal safe charging of the battery. It also has the capability to act as a fuel gauge for the battery estimating the remaining charge level. The battery management system functions as a maximum

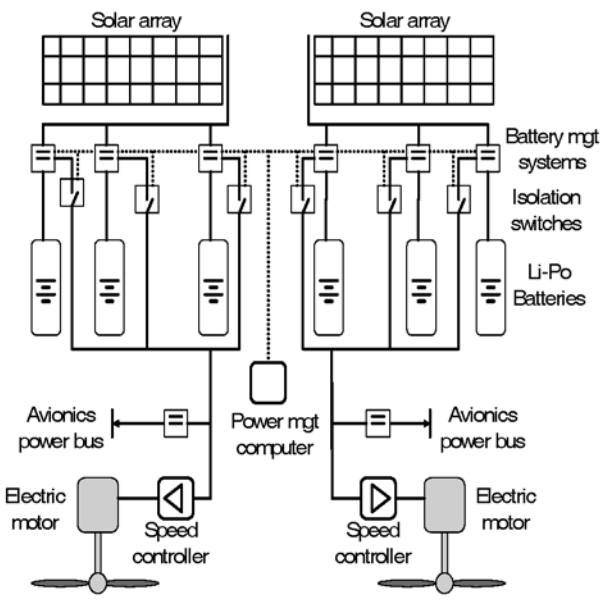


Figure 6: Solar powered UAV power system.

power point tracker ensuring that the maximum electrical power can be extracted from the solar array. The battery management systems are connected to the power management computer by means of a Control Area Network (CAN) bus.

4.3 Isolation Switches

Isolation switches are used to isolate the batteries from the electric power busses of the UAV. During the charge cycle only one battery on the left and one battery on the right are connected to the UAV electric system while the other four batteries are being charged from the solar array.

4.4 Energy Storage Devices

The purpose of the energy storage devices is to store the excess electrical energy collected by the solar arrays during day time and make it available during night time. Lithium polymer batteries were selected as energy storage devices primarily due to their high energy capacity to weight ratio. The selected batteries were connected in a 6S2P configuration producing a maximum voltage of 25.2 V when fully charged and 18 V when discharged. For a 7400 mAh capacity battery the weight is 1.117 kg giving a total battery weight for the six batteries of 6.7 kg. The total energy storage capacity of the batteries is 3.53 MJ.

4.5 Power Management Computer

The power management computer acts as the interface between the telemetry system and the battery management systems, selecting the active drive batteries while isolating the batteries being charged. State of charge information from each battery is obtained via the CAN bus network and then relayed to the telemetry system for monitoring by the ground station.

4.6 Electric Motors

Brushless DC motors of the out-runner type are selected as the electric propulsion motors for the UAV. These motors offer high torque at high efficiencies of around 88%. Since the efficiency of these motors increase at higher revolutions the choice was made to use two smaller 500 W motors running at higher revolutions. For initial take-off and climb to altitude the power requirements for a 12 kg UAV is estimated at 1 kW. Once level flight conditions are achieved the power consumption is drastically reduced to an estimated 80 W which results in each motor running at 8% of its rated power. The motors selected are the AXI3140/60 brushless DC motors driving 14" propellers. Speed control of each motor is achieved by standard radio control brushless motor speed controllers receiving control inputs from the onboard radio control receiver.

4.7 Avionics

To enable the UAV for autonomous flight an autopilot from ABAT is installed as part of the avionics suite [12]. The autopilot is connected to the telemetry system from were it can receive over-riding commands from the ground control station. Back-up control is facilitated using a spread spectrum radio transmitter operating in the 2.4 GHz frequency band. The controllability of the UAV in case of an onboard radio receiver failure is further enhanced by using two onboard radio receivers. The one radio receiver is connected to the normal control servos of the UAV while the other receiver is connected to the flap servos. In this manner should the primary radio receiver fail at least a measure of roll control can be achieved by operating the flaps in a differential manner.

5. SOLAR FLIGHT SIMULATION

For the validation of long endurance low altitude solar flight a simulation was realized using MATLAB Simulink. The schematic of the Simulink simulation model is presented in Figure 7. The simulation model includes the radiation model which is parameterized by the geographical position and the time of the year. A facility is provided whereby the flight start time can be entered as an offset to the simulation time. The model of the solar panel takes into account the solar array size and efficiency. The power generated by the solar array is then made available to the battery model. An aircraft model was implemented which computes the required power for level flight. Electrical power required by the avionics is added to the power required for level flight to obtain the total power requirements for the UAV.

Figure 8 shows the power progression graph for the power flow of the UAV. The flight was started at 18:00 in the evening on summer solstice with a fully charged

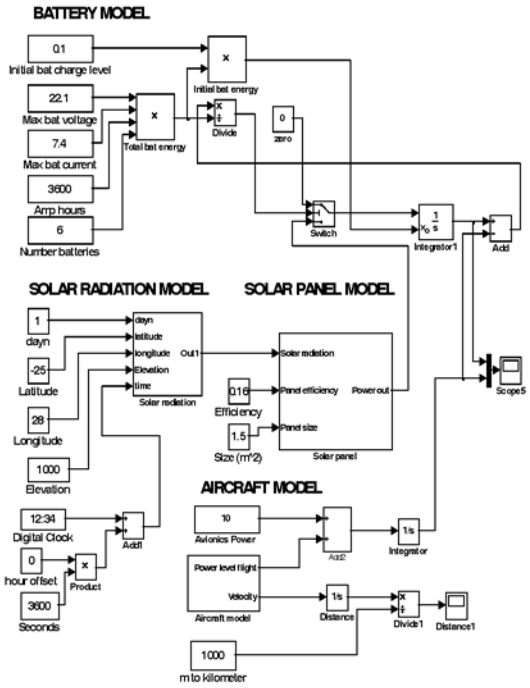


Figure 7: Schematic of the solar flight simulation model using Matlab Simulink.

battery. Flight duration of 48 hours is simulated in order to complete the required distance of 1400 km. The solar power is the power collected by the solar array. UAV power includes the power required for maintaining level flight and the power required by the onboard avionics.

The power flow graph depicted in Figure 8 does not give a clear understanding of how much power should be collected by the solar array and what capacity energy store is required. Better insight into the power balance requirements can be obtained by plotting the energy flow graph of the UAV as in Figure 9.

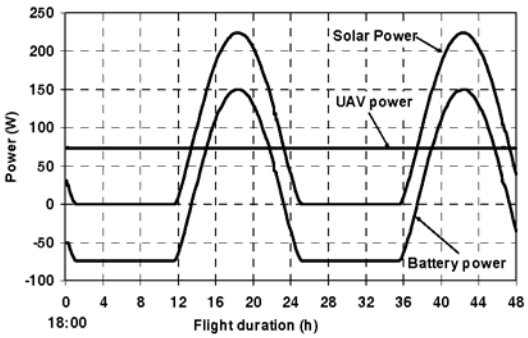


Figure 8: Power flow for a long endurance solar powered UAV requiring 74 W for maintaining level flight. A solar array of 1.2 m^2 with efficiency of 16% is collecting the solar radiation.

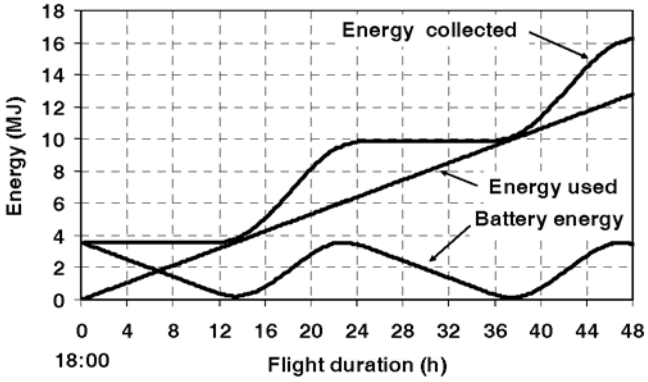


Figure 9: Energy flow for a long endurance solar powered UAV for a flight duration of 48 hours.

From Figure 9 it can be seen that the energy collected by the solar array always exceeds the energy used by the UAV fulfilling the requirement of Equation 2. The contribution of the fully charged battery can be seen as the flight starts at 18:00 in the evening. The initial battery charge is nearly depleted at 06:30 in the following morning from where charging resumes to full capacity at 17:00 in the afternoon.

As the available light diminishes towards the evening the battery starts to discharge to supplement the dwindling available solar energy. At about 06:00 in the morning of the second day the battery is again nearly discharged when the following charge cycle begins. From the battery energy graph shown in Figure 9 the required battery capacity for sustained flight is obtained as 3.51 MJ.

6. CONCLUSION

From the work presented it can be concluded that the primary consideration in the design of low altitude long endurance UAVs is the available solar energy which can be collected by the photovoltaic array mounted on the wing surface of the UAV. The available solar energy is very strongly dependant on the time of year and the clearness of the atmosphere. The second consideration in the design of low altitude long endurance UAVs is the design of the wing which is the primary aerodynamic surface. The wing should be designed for high lift and low drag at low speed to ensure the best power factor giving the minimum thrust requirement. During operation the UAV must be trimmed for optimal speed ensuring maximum endurance. The capacity of the energy store can be determined by the power and energy requirements of the UAV. By taking these considerations into account it should be possible to design and build a low altitude long endurance solar powered UAV capable of demonstrating sustained flight. The next phase in the development of the LALE UAV would be the construction of the UAV and the execution of long endurance piloted flight tests before reverting to full autonomous operation.

7. REFERENCES

[1] Panasonic World Solar Challenge, Across Australia, Darwin to Adelaide, 21-28 October 2007, http://www.wsc.org.au

[2] K. Kowalenko, UAV Technology to Aid in Disasters, The Institute, IEEE, September 2001.

[3] K. C. Reinhardt, T. R. Lamp, J. W. Geis, A. J. Colozza, Solar Powered Unmanned Aerial Vehicles, Energy Conversion Engineering Conference, 1996, IECEC 96, Proceedings of the 31st Intersociety, Volume 1, pages 41 to 46.

[4] A. Colozza, Fly Like a Bird, IEEE Spectrum, pages 32 to 37, May 2007

[5] A. Cocconi, AC Propulsion's Solar Electric Powered UAV, http://www.acpropulsion.com

[6] I. Genuth, Solar Powered UAV to Set the World Endurance Record, The Future of Things, 9 March 2007, http://www.TFOT.info

[7] R. E. Bird, R. L. Halstrom, A Simplified Clear Sky Model for Direct and Diffuse Insolation on Horizontal Surfaces, Solar Energy Research Institute, SERI/TR-642-761, Feb 1981.

[8] F. Kasten, The Linke Turbidity Factor Based on Improved Values of the Integral Rayleigh Optical Thickness, Solar Energy 56: 239-44.

[9] M. Simons, Model Aircraft Aerodynamics, Special Interest Model Books, Fourth Edition, 2002.

[10] Athena Vortex Lattice v3.26, H. Youngren, M. Drela, MIT

[11] X-Plane v 8.5, Laminar Research, http://www.x-plane.com

[12] ABAT, http://www.abat.co.za

NEGATIVE EFFECTS OF ENERGY-SAVING, NON-LINEAR LOADS ON LV SYSTEMS: CAUSES AND RECOMMENDATIONS

R. Herman*, C.T. Gaunt* and G.S. Raubenheimer**

** Department of Electrical Engineering, University of Cape Town, Private Bag, Rondebosch, South Africa*
*** Dept. of Electrical and Electronic Engineering, University of Stellenbosch, Private Bag X1, Matieland, 7602, South Africa*

Abstract: Pressure is being exerted on utilities and electrical energy customers to reduce energy consumption. There are two concerns: the effect that increasing greenhouse gasses have on global warming and the dwindling non-renewable resources currently used for the production of electrical power. The use of modern energy-efficient appliances such as Compact Fluorescent Lamps and micro-wave ovens is a widely promoted solution. These types of devices are recommended not only in affluent areas but also in developing countries (such as South Africa) where the use of computers and television sets are also being encouraged as devices for increasing the level of information and education. However, the use of these devices can introduce significant problems. This paper reviews a range of published material on the various aspects of this topic. It also includes the results from experimental measurements. Finally, the paper discusses some recommendations that are intended to bridge the gap between theoretical analyses and the practical implementation of mitigating steps, as they apply to low voltage distribution systems.

Key words:
Non-linear loads; harmonics; measurement and verification; wiring installations

1. INTRODUCTION

There is global concern over the current levels of greenhouse gas emissions, to which the generation of electricity from fossil fuels is a major contributor. Efforts to reduce emissions fall into two main categories – (a) make more use of renewable energy sources such as wind power, photovoltaic power and power from bio-mass, and (b) reduce energy consumption by using more efficient appliances. The second of these mitigation categories invariably employs devices with non-linear electrical characteristics. In this text these loads are referred to as Non-Linear Loads (NLLs). Their use is increasingly evident in the domestic as well as the commercial sectors.

In South Africa, the social, political and economic impact of the extensive electrification programme has been significant during the past 10 years. Yet, approximately 3 million households still need to be electrified, mostly in rural and so-called 'deep-rural' areas. These customers largely fall into the poorest class where unemployment is high and the level of education is low. In addition, the peak demand is growing and the national utility, Eskom, is facing a serious capacity problem, which is expected to become critical in 2007.

More efficient use of the available power is essential. A seeming logical option is to encourage the use of energy-efficient lighting in the form of Compact Fluorescent Lamps (CFLs) and the use of micro-wave cooking. Traditional food preparation and relatively high capital costs would influence the effectiveness of the latter. Demand side intervention may also be facilitated through education and training in the use of electrical energy. The most effective way of achieving this is via the television network, which obviously requires the use of television sets. These devices also have the disadvantage of distorting the load current. It should be understood that most of the rural customers cannot afford to purchase and operate high-energy appliances such as water heaters, electric ranges and air-conditioners. Consequently, the composite rural customer load may have a lower proportion of linear loads than the higher income groups in urban locations.

The occurrence of non-linear loads is equally prevalent in the commercial sector, particularly in the urban setting. Modern commercial loads are comprised of ever increasing numbers of computers, fluorescent light fittings (including CFLs) and air-conditioning units. Variable speed drives are now being used to operate the air-conditioning equipment in, ostensibly, the most efficient manner.

This paper attempts to highlight the negative effects of modern non-linear loads on low voltage (LV) systems based on the calculations and measured results derived by the authors and previously published literature. It will show what steps could be taken to bridge the gap between knowledge and technology, particularly in the South African context.

The problem has practical implications. These essentially arise from the gap between the theoretical analyses of the

phenomena and a practical guideline to combat the problems. Operators and contractors at the LV distribution level require binding, lucid directives that will ensure the lowest penalty to the system and the customer.

So, while it is desirable to use 'energy-efficient' appliances to reduce energy and maximum demand, the electrical characteristics of these devices inevitably lead to the generation of unwanted harmonic currents. A large volume of literature has been published during the past two decades on the theoretical assessment of these harmonic generating devices. Some of the findings from the published articles will be reviewed in this paper. Results obtained from experimentation are also discussed. From these results recommendations to remedy the problems are proposed and discussed.

2. LITERATURE SURVEY

The material reviewed in this section is illustrative rather than exhaustive. It serves to identify the main issues relating to the generation of current harmonics by modern NLLs and the consequent impact on the LV network to which they are connected.

2.1 *Power definitions and relationships*

In current distribution practice almost all the calculations used in the sizing of electrical equipment are based on phasor quantities. These fundamental relationships have come under greater scrutiny with the increase in the incidence of NLLs. This topic is discussed at some length in the paper prepared by the *IEEE Working Group on Non-sinusoidal Situations* [1].

Under the assumption of ideal sinusoidal conditions the traditional approach to power relationships in a-c networks yields the following:

$$\begin{aligned} v &= \sqrt{2}.V\sin(\omega t) \\ i &= \sqrt{2}.I\sin(\omega t - \varphi) \end{aligned} \tag{1}$$

where, v and i are instantaneous values, and V and I are RMS values, of voltage and current. The relationships in Equation1 assume constant angular velocity,ω [rad/sec] of the voltage and current phasors. The quantity φ represents the phase angle between the voltage and current phasors.

Then:

$$\begin{aligned} P &= VI\cos\varphi \\ S &= VI \\ Q &= VI\sin\varphi \\ S &= \sqrt{P^2 + Q^2} \end{aligned} \tag{2}$$

Where, P is the active power [W], S the apparent power [VA] and Q the non-active (often called reactive) power [VAr]. The power factor is then given alternatively as:

$$\begin{aligned} PF &= \cos\theta \\ PF &= \frac{P}{S} \end{aligned} \tag{3}$$

The quantity P is the rate at which energy is transferred to the load. It is further to be noted that energy in electrical terms can only be measured by an energy or revenue meter in energy units – Joule-sec, Watt-hours, MWh etc. In simple terms, energy is the quantity that does work and is uniquely associated with the active power, P.

Equations 1 to 3 lose their validity when either, or both, of the voltage and current waveforms depart from their sinusoidal shape. Under these conditions one cannot refer to a phase angle, because ω and φ are no longer constant. Hence, power factor varies in time throughout the fundamental period, T. We then have to use expressions with greater generality, such as those given by Sharon [2]:

$$\begin{aligned} P &= \frac{1}{T}\int_0^T vi.dt \\ S &= \frac{1}{T}\sqrt{\int_0^T v^2 dt.\int_0^T i^2 dt} \end{aligned} \tag{4}$$

The power factor, the ratio of power to apparent power, is still an indicator of the effectiveness of the delivery of energy.

Non-sinusoidal, periodic voltages or currents may be expressed as a sum of sinusoidal harmonics, for example:

$$i = i_1 + i_2 + i_3 + ... + i_n \tag{5}$$

$$i_1 = I_1 sin(\omega t\text{-}\varphi_1); \quad i_n = I_n\, sin(n.\omega t\text{-}\varphi_n) \;\; etc.$$

Where the subscripts refer to the harmonic order and I_1 to I_n are the magnitudes of the harmonic currents. To facilitate the algebra we can simplify Equations 4 and 5 by letting $\theta = \omega t$, $dt = 1/\omega.d\theta$ and $T = 2\pi$.

If the LV distribution is 'stiff' enough (i.e. it has low Thevinin impedance), we can assume that the supply voltage at the point of a single-phase, non-linear load (NLL) is sinusoidal and free of harmonics. The current drawn by the NLL will be given by an expression similar to Equation 5 above. Then upon substitution into Equation 4 we obtain in general terms for any number of harmonic current components:

$$P = \frac{V_1 I_1 \cos\varphi}{2}$$
$$S = \frac{V_1 \left(\sum_{h=1}^{n} I_h^2\right)^{\frac{1}{2}}}{2} \tag{6}$$

Using the relationship *P/S* for the power factor, PF, we obtain:

$$PF = P/S = \cos\varphi . \frac{I_1}{\left(\sum_{h=1}^{n} I_h^2\right)^{\frac{1}{2}}} \tag{7}$$

It is clear from Equation 7 that the power factor reduces as the size and number of the harmonic components increases.

To illustrate the effects on power factor and thermal losses in conductors, consider a NLL with current component magnitudes: I_1 = 10A; I_3 = 7A; I_5 = 4A and I_7 = 4A. Take the phase angle of the fundamental component as 26^0. Assume further that the NLL is supplied from a 230V sinusoidal voltage source and that the total conductor resistance to the load is 0.5 ohms. The results given in Table I are then obtained.

Table I: Sample Calculations

Active Power	[W]	2071
Apparent Power	[VA]	3094
Cos(φ)		0.9
Power Factor	P/S	0.67
Conductor Losses	[W]	45.25
Losses due to I_1	[W]	25

The results suggest that under the given conditions, the power factor has worsened from 0.9 to 0.67 and the conductor losses have increased by 80%. Conductor losses are likely to be even higher because, due to 'skin-effect', the effective resistance of the conductor increases as the frequency of the current increases. The increase in resistance for the higher order harmonics will increase the estimated 'copper-losses'.

2.2 *Quality of supply*

Quality of supply (QOS) embraces a wide variety of aspects and these will not be considered in detail in this paper. However, one of the definitions, the total harmonic distortion of the current (THD_I), is helpful in describing and analysing the effects of NLLs:

$$I_H = \sqrt{\sum_{h=2}^{\infty} I_h^2}$$
$$THD_I = \frac{I_H}{I_1} \times 100\% \tag{8}$$

In these expressions I_h is the RMS value of the hth harmonic current component and I_1 is the RMS value of the fundamental current component. Typical measured values for domestic and commercial appliances are discussed in Section 3.

2.3 *Compact fluorescent lamps (CFLs)*

During the past two decades, various authors have reported on the current harmonic production of CFLs and its deleterious effects on distribution systems. For example, Mielczarski has drawn attention to a variety of side effects resulting from the use of CFLs [3]. These include accelerated equipment aging and interference with communication systems. Pileggi et al. performed simulations to investigate the effect of CFL loads on voltage distortion on a typical distribution feeder [4]. Within the context of the current trend to shift more generation into the renewable sector, it would be well to take note of the observations documented by Korovesis et al. [5]. Their paper investigates the influence of a large-scale installation of CFLs on line voltage distortion of a weak network supplied by a photovoltaic station.

2.4 *Personal computers (PCs)*

The personal computer (PC) made its appearance in the early 1980's. Their proliferation since then has been exponential and they presently represent a significant load component, particularly in commercial and academic institutions. These devices are fed from switch-mode power supplies and inherently draw distorted load current. A number of researchers have investigated the generation of this current harmonic distortion and the resultant impact on the networks. Moore [6] reports on his investigation into the influence of PC processing modes on line current harmonics. In an extensive survey of more than 370 PCs his research shows that hard-drive access causes the highest level of 3rd and 5th harmonic load currents. He notes that the degree of harmonic attenuation depends on the number of PCs connected to the same supply. But the diversity effects due to variation in loading level showed an increase in harmonic generation with line current RMS level. Diversity effects are also investigated by Grady et al. [7]. Aintablian and Hill [8] draw attention to the increase in neutral current in three-phase, 4-wire systems that arises from loads with triplen harmonics (3rd, 9th etc.). This phenomenon will be addressed in more detail later in this paper.

2.5 *Switch-mode power supplies*

Switch-mode power supplies are not only confined to PCs, but may be found in uninterruptible power supplies (UPS), printers, photocopiers, battery chargers etc. All these devices draw highly distorted load current. Further reports on the harmonics of switch-mode power supplies are found in literature [8, 9].

2.6 *Adjustable speed drives (ASDs)*

The speed of induction motors may be controlled by varying the supply frequency. This is commonly

achieved by using pulse-width modulation (PWM), a process that leads to the generation of odd current harmonics, particularly the 5th, 7th, 11th and 13th. ASD applications are common in HVAC systems (heating, ventilating and air-conditioning), industrial processes and uninterruptible power supplies (UPSs). Modern energy-efficient equipment often use ASDs. Typical of the research done in this field are the papers by Dell'Aqila [10] and Capasso [11].

2.7 *Energy metering*

Customers have the right to accurate metering of their electrical consumption. However, the current distortion introduced by some of the energy-saving loads can, under certain conditions, cause the energy metering to be erroneous. Some utilities apply a tariff penalty for so-called 'reactive energy'; exactly what the term means and how it is to be measured when the current waveform is distorted is not defined. A number of research papers have been published on the theme of energy metering loads with distorted currents. Czarnecki comments on the effect of distorted current waveforms on active power flow and energy accounts [12]. He concludes that customers with polluted loads, when billed for energy measured as the integral of active power, cause a loss of revenue to the utility. Arseneau [13] supports this view and examines the performance of revenue meters based on existing (2004) definitions in comparison with one that measures the energy using only the fundamental waveform. In another paper [14] he examines the IEEE Standard 1459-2000 for revenue meters and concludes that it is inadequate for the measurement of NLLs.

2.8 *Demand side measurement and verification*

Related to the previous section is the measurement and verification of demand side management (DSM) interventions. There are many programmes worldwide that encourage the use of DSM strategies. Customers often receive financial incentives from the utilities, based on the measured performance of the DSM initiatives. Common programmes often include the replacement of incandescent lighting with CFLs and retrofitting of HVAC systems with compressors using ASDs. As mentioned, these devices draw distorted currents. Therefore, such post-installation loads will have a high component of polluted load current and, as indicated above, the measurement and verification of these interventions are subject to errors. Traditional induction disc meters under-estimate the revenue payable to the utility. These issues are mostly ignored in related operational standards, such as the American document: *M & V Guidelines: Measurement and Verification for Federal Energy Projects, Version 2.2* [15]. Practice in South Africa uses this document as a basic guide but also ignores the harmonic distortion of load currents.

3. EXPERIMENTAL RESULTS

The material presented and discussed in this section was derived from experimental measurement. These measurements show the extent of current distortion in the loads commonly associated with energy saving exercises.

The table in each of the following sub-sections 3.1 to 3.4 presents the RMS values of voltage and load current, the apparent power as recorded on a digital power analyser, the total harmonic distortion (THD), and the power factor obtained from the power analyser. Typical waveforms and the harmonic content as a percentage of the fundamental are illustrated graphically. In all cases, the 3rd harmonic content of the waveform is significant.

3.1 *Compact fluorescent lamps (CFLs)*

Measurements in this section refer to a 21W Compact Fluorescent Lamp.

Table II: CFL Measurements

V rms	I rms	S VA	THD %	PF
216	0.142	30.672	106.4	0.58

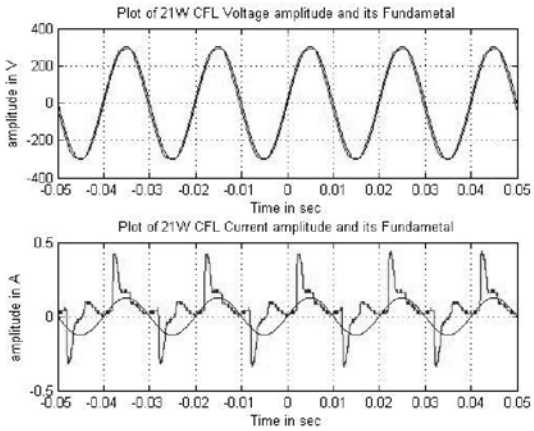


Figure 1: CFL Voltage and Current Waveforms

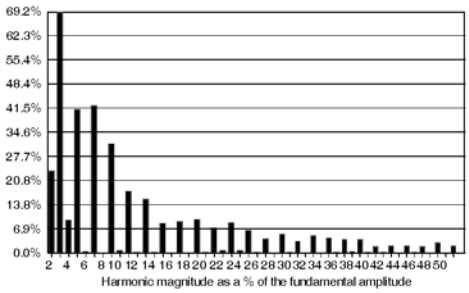


Figure 2: CFL Current Harmonics

3.2 *Personal computers (PCs)*

Table III PC Measurements

V rms	I rms	S VA	THD %	PF
225	1.26	283.5	121.17	0.63

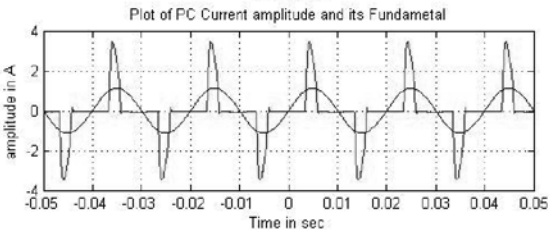


Figure 3: PC Current Waveform

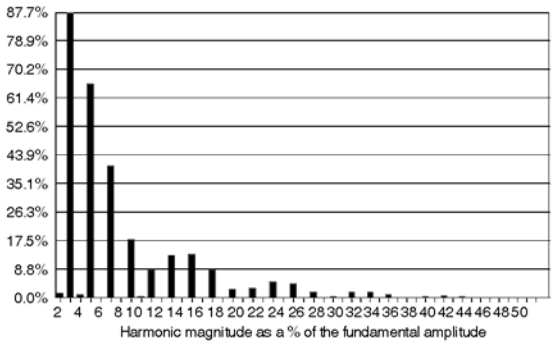


Figure 4: PC Current Harmonics

3.3 Television sets (TVs)

Table IV: TV Measurements

V rms	I rms	S VA	THD %	PF
225	0.563	126.675	155.47	0.441

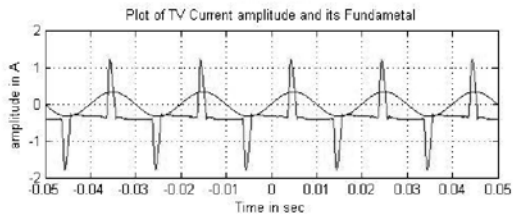


Figure 5: TV Current Waveform

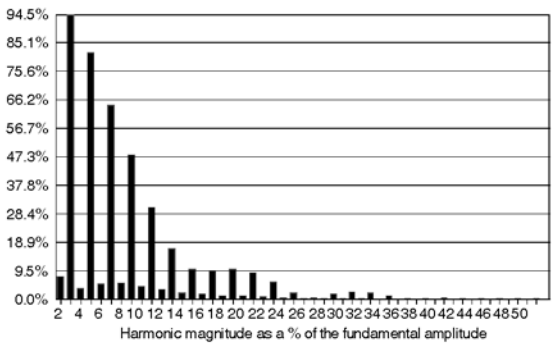


Figure 6: TV Current Harmonics

3.4 Microwave oven

Table V: Microwave Oven Measurements

V rms	I rms	S VA	THD %	PF
227	4.98	1130.46	38.47	0.871

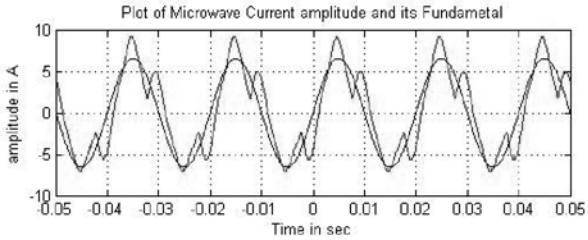


Figure 7: Microwave Current Waveform

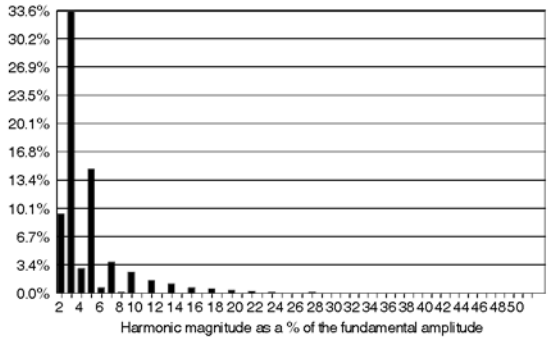


Figure 8: Microwave Oven Current Harmonics

3.5 Energy metering

The energy consumption of various appliances was measured with a traditional Ferraris type energy meter and the results compared with a true-reading digital energy meter. The percentage errors are summarised in Table VI.

Table IV: Energy Measurement Errors

Appliance	Error %
9W CFL	24.82
PC	7.59
TV	25.62
Microwave	3.9

The measurements indicate a loss of revenue to the utility for these loads when they are measured with conventional kWh metering.

3.6 Effect of neutral current from PC loads

The use of personal computers (PCs) is increasing at phenomenal rate. Because of their value as aids to education, they are often foremost in donations to the poor in developing countries, such as to schools in these communities. The impact on the loading of a network feeding a large number of computers presents real problems to the distribution systems.

The load illustrated in Figures 8 and 9 was recorded at a computer lab (larger than one in a typical school) in which a large number of PCs were supplied.

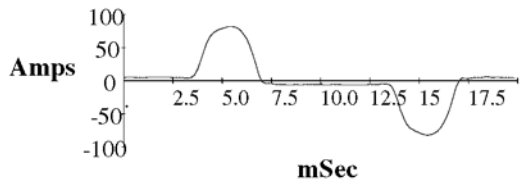


Figure 8: Current Waveform Blue-phase

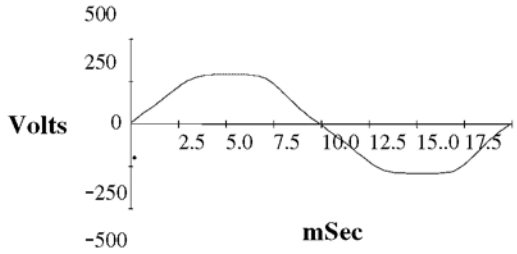


Figure 9: Voltage Waveform Blue-phase

The 3rd harmonic current was 27% of the fundamental. The large distortion in current causes voltage drops in the phase conductors feeding the PCs that result in a serious distortion of the voltage supply waveform, as shown in Figure 9. In a three-phase supply network the largest concern is the summation of the co-phasor triplen current components (3rd and 9th harmonics) in the neutral conductor. In measured cases in our study some neutral currents were found to be approximately 25% larger than the phase current.

4. IMPLICATIONS

Some of the many implications of NLLs are summarised here to indicate the remedial actions required.

Power factor definition and evaluation: In traditional network calculations the power factor is usually conceived and referred to as the cosine of the angle between the voltage and current phasors. It was shown in Section 2.1 that this is not valid when the load current has harmonics. Definitions of real and apparent power must be modified under these conditions.

Conductor heating losses: The sample calculations in Section 2.1 show how harmonics can increase ohmic losses in conductors for a given transfer of active power. These increased losses are largely ignored in network calculations.

Non-active (reactive) power compensation: Capacitors are often used in to correct power factor in networks with low 'lagging' power factors. The presence of NLLs has two negative implications**:** the higher frequencies can damage the capacitors and low power factor that varies instantaneously within each cycle is difficult to compensate with static filters.

Load modelling: Traditional Z, I, P models may no longer be valid when the loads are comprised of a significant NLL component, because the load characteristics change even within the normal range of voltage variation.

Load data sampling: Load data is usually monitored and stored with the use of digital logging equipment. Load parameters are sampled at a fixed sample rate. The Nyquist criterion is usually applied in these cases, which requires sampling at twice the fundamental frequency. Obviously, when measuring currents (and powers) with high harmonic content a faster sampling rate is required, e.g. 300 Hz for the 3rd harmonic of a 50 Hz system.

Network parameter calculations and models: Network parameters are traditionally calculated as resistance, inductive- and capacitive-reactance at fundamental frequency. Inductive reactance is directly proportional and capacitive reactance indirectly proportional to the frequency. The network parameters might have to be modified to include the effects of the third harmonics (and possibly the higher order ones as well) [16].

Voltage distortion: As shown in the example of the computer laboratory the voltage drop with its distorted waveform can modify the supply voltage, typically giving it a flattened peak. This could lead to quality of supply problems in sensitive areas. In domestic installations, where NLLs comprise a significant component of the total load, and where induction motors (e.g. refrigerator motors) are connected, the motors could experience over-heating.

Errors in revenue metering: The metering error discussed in Section 3.5 could lead to a significant loss of revenue in the case of NLLs. This could include domestic and commercial loads. Larger industrial loads using ASDs and power electronic equipment would have similar effects.

DSM intervention measurement: As discussed in Section 2.8, the measurement of pre- and post installation demand becomes suspect in the presence of NLLs. Erroneous measurement during verification will misrepresent the effects of DSM interventions that focus on replacement with energy-efficient appliances.

Triplen currents in the neutral: One of the most severe implications occurs in four-wire three-phase systems, common in distribution networks and wiring installations. Distorted currents with 3rd and 9th harmonic components give rise to triplen currents that are co-phasor in the neutral, as reported in Section 3.6. These can cause overheating and damage in a neutral conductor, which is often smaller in cross-sectional area, and therefore of higher resistance than the phase conductors. This could lead to degradation of insulation, short circuit currents and fire damage. These conditions are not covered in standard wiring regulations.

Transformer over-heating: The presence of harmonic components increases the eddy current losses in the core of a supply transformer. Triplen harmonic currents circulate in the delta windings of transformers, increasing conductor losses. At the higher frequency of the harmonic components of current, the leakage flux is deformed leading to an increase in losses. These increased losses raise the temperature of a transformer supplying significant NLLs, increasing the insulation degradation and reducing the lifetime, which may require de-rating of a transformer.

Auxiliary relays: Harmonic components can exceed the withstand capacity of small switching devices. For example, a photo-sensitive switch rated for loads up to 100 W failed in a few days when used to control three CFLs with a total power of only half the apparent rating of the switch. The switch, shown in Figure 10, over-heated and in the 'right' circumstances could have initiated a fire.

5. RECOMMENDATIONS

5.1 Harmonic mitigation

The obvious mitigation action would be to eliminate the generation of harmonics at the load. This would require cooperation from the manufacturers of appliances to ensure the suppression of harmonics, and national and international standards to guarantee compliance. Research will be needed to establish suitable limits for THD and 3rd harmonic content, which, it is proposed, would be displayed on the nameplates of all electrical equipment. A specification for NLLs would need to be harmonised with quality of supply specifications, such as NRS 048. Such an approach to mitigating NLL effects would require time (years) for implementation and would not eliminate the problem from existing equipment.

5.2 Wiring code

Existing wiring codes in South Africa do not adequately address the issues related to the use of NLLs. Three aspects need attention:

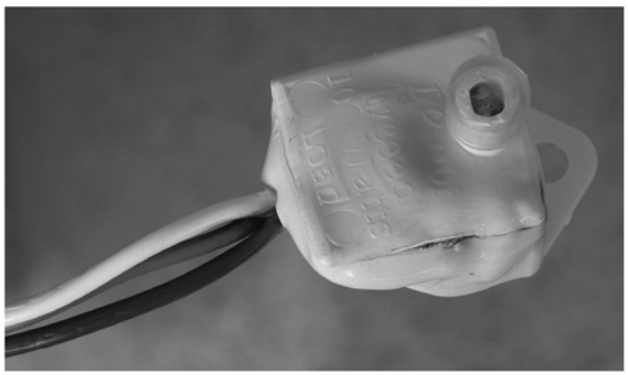


Figure 10: Switch Overheated by CFL Load

- The wiring code should give guidance in assessing the level of NLLs. For example, contractors would need to know how to assess a circuit when a significant proportion of the load consists of PCs, CFLs, ASDs or other NLLs.

- Triplen components cause large current magnitudes in the neutral of a 4-wire, three-phase system. Therefore the neutral conductor of all 4-wire circuits to NLLs should have an equivalent thermal rating at least equal to, and preferably 1.5 times higher than, the rated phase current.

- It might be difficult and costly to replace neutral wiring in existing installations. An alternative would be to employ a protection device that trips all three phases when (a) the phase current limit is reached and/or (b) when the neutral current exceeds the conductor current rating. In such a case the neutral conductor is not opened by the protection device. This may be accomplished practically by using four mechanically ganged MCBs and a short circuit across auxiliary contacts of the 4th (neutral) MCB.

5.3 Appliance and device labelling

In addition to the proposals for reducing harmonic generation discussed in Section 5.1, the following requirements for appliance and device labelling should be introduced as quickly as possible:

- Relays and switches should be specified to operate with NLLs, or should otherwise be clearly identified as being suitable only for resistive loads.

- Appliance ratings and labels should be expressed in both real and apparent power to reduce the risk that circuits will be under-sized because of inadequate allowance for the low power factor of appliances and equipment incorrectly assumed to be resistive.

5.4 DSM measurement & verification

It is recommended that DSM measurement and verification guidelines be thoroughly examined and rewritten. The areas that require particular attention are:

- *Electrical power definitions in the presence of non-sinusoidal voltages and/or currents:* The definitions must have universal acceptance. These definitions must include active, non-active and apparent power as well as power factor.
- *Measurement of active, non-active and apparent power and power factor:* Specifications must include standard revenue metering as well as logging equipment for verification.
- *Sampling interval for measurement equipment:* It is assumed that electrical power measurements will in future be done using static electronic functional

circuit elements (chips). The sampling rate for the measuring devices must be fast enough to ensure accuracy when the waveforms are non-sinusoidal.

5.5 Tariffs

The comments in Section 5.3 largely refer to DSM interventions but may also be applied to tariff assessment and development. It is important to structure tariffs that penalise customers whose load currents are distorted in such a way that verification and correction are possible. For this purpose both the true power factor (using a revised definition from Section 5.2) and the THD of the load current should be used.

5.6 Replacement of Ferraris meters

It is clear from the discussion that the conventional Ferraris meter is inadequate for measuring energy when the current is distorted. These should be replaced with static meters that measure true power.

5.7 Transformer rating

Transformer over-heating may be prevented by de-rating the transformer by a suitable amount. Other precautions include specifying and designing the transformer to limit the increased losses caused by NLLs.

6. CONCLUSION

The widespread adoption of appliances and equipment with non-linear electrical characteristics, as a way of increasing energy efficiency, has the potential to significantly and negatively affect supply systems. The use of appliances and control systems (ASDs) that distort the fundamental sine wave of the power supply system affects not only the quality of supply parameters, but has practical implications for the design and operation of supply systems. Urgent attention must be given to developing new definitions, specifications and practical procedures; to labelling appliances appropriately; to metering supplies accurately; and to maintaining the integrity and safety of electrical installations.

ACKNOWLEDGEMENTS

Part of this research has been supported by funding from the Department of Trade and Industry's THRIP and Eskom's TESP, which is gratefully acknowledged.

7. REFERENCES

[1] IEEE Working Group on Non-Sinusoidal Situations: Effects on Meter Performance and definitions of power: "Practical Definitions for Powers in Systems with Non-sinusoidal Waveforms and Unbalanced Loads: A discussion", *IEEE Trans Power Delivery, Vol. 11 January 1996, pp 79-101.*

[2] Sharon, D., "Power factor definitions and power transfer quality in a non-sinusoidal situation", *IEEE Trans., Instrumentation and measurements*, 1996, Vol 45 (3), 728 – 733.

[3] Mielczarski, W, Michalik, G, Lawrence, W.B. and Gabryjelski, Z, "Side effects of energy saving lamps", *IEEE: 8th Int. Conference Harmonics and Quality of Power, 1998, pp 1211 – 1217.*

[4] Pileggi, D.J, Gulachenski, E.M, Root, C.E, Gentile, T.J and Emanuel, A.E., "The effect of modern compact fluorescent lights on voltage distortion", *IEEE Trans Power Delivery, Vol. 8 July 1993, pp 1451 – 1459.*

[5] Korovesis, P.N, Vokas, G.A Gonos, I.F and Topalis, F.V, "Influence of large-scale installation of energy saving lamps on the line voltage distortion of a weak network supplied by photovoltaic station", *IEEE Trans Power Delivery, Vol. 19 Oct 2004, pp 1787 – 1793.*

[6] Moore, P.J and Portugués, "The influence of personal computing processing modes on line currents", *IEEE Trans Power Delivery, Vol. 18 October 2003, pp 1363 – 1368.*

[7] Grady, W.M, Mansoor, A, Fuchs, E.F, Verde, P and Doyle, M, "Estimating the net harmonic currents produced by selected distributed single-phase loads: Computers, Televisions and Incandescent light dimmers", IEEE Power Engineering Society Winter Meeting, 2002, pp 1090 – 1094.

[8] Plow, R, "Effects of switch mode power supply harmonic currents on building power systems and stand-by engine alternators", Telecommunications Energy Conference, INTELEC '94, 1994, pp 665 – 668.

[9] Key, T.S and Jih-Sheng, L, "Costs and benefits of harmonic current reduction for switch-mode power supplies in a commercial office building", *IEEE Trans Industry Applications, Vol 32, Sept/Oct, 1996, pp 1017 – 1025.*

[10] Dell'Aquila, A, Lassandro, A and Zanchetta, P, "Modelling of line side harmonic currents produced by variable speed induction motors", *IEEE Trans Energy Conversion, Vol 13, Sept, 1998, pp 263 – 269.*

[11] Capasso, A.; Lamedica, R.; Prudenzi, A "Estimation of net harmonic currents due to dispersed non-linear loads within residential areas", IEEE: 8th International Conference on Harmonics and Quality of Power, 1998, pp 700 – 705.

[12] Czarnecki, L.S, "Comments on active power and energy accounts in electrical systems with non-sinusoidal waveforms and asymmetry", *IEEE Trans Power Delivery, Vol. 11 Jul 1996, pp 1244 – 1250.*

[13] Arseneau, R and Hughes M.B, "Selecting revenue meters for harmonic producing loads", 11th Conference on Harmonics and Quality of Power, 2004, pp 227 – 231.

[14] Arseneau, R, "Application of IEEE standard 1459-2000 for revenue meters", IEEE - PES General Meeting, July 2003, pp 87 – 91.

[15] M & V Guidelines: Measurement and Verification for Federal Energy Projects, Version 2.2, pp 111 – 144.

[16] El-Saadany E.F. and Salama M.M.A, "Reduction of the net harmonic current produced by single-phase non-linear loads due to attention and diversity effects", *Electrical Power and Energy Systems, Vol. 20, pp 259 – 268, 1998.*