MSC ENGINEERING DISSERTATION

THE USE OF STEADY-STATE LEVEL COMBINATIONS AND SIGNAL EVENT EDGE CORRELATIONS IN THE DISAGGREGATION OF TOTAL POWER MEASUREMENTS

Joseph J. Penn

A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering.

Johannesburg, 2015

DECLARATION

I declare that this dissertation is my own unaided work except where specifically acknowledged. It is being submitted for a Degree of Master of Science to the University of the Witwatersrand, Johannesburg. It has not previously been submitted for any degree or examination to any other university.

Joseph J. Penn

Date

ABSTRACT

The work presented extends and contributes to research in Non-Intrusive Load Monitoring (NILM), focussing on steady-state and transient power measurement disaggregation techniques for circuits containing household appliances. Although previous work in this area has produced and evaluated a wide range of NILM approaches, much of it has involved the use of datasets captured from real-world household implementations. In such cases, the lack of accurate ground truth data makes it difficult to assess disaggregation techniques. In the research presented, three NILM techniques are comparatively evaluated using measurements from typical household appliances assembled within a laboratory environment, where accurate ground truth data could be compiled to complement the measurements. This allows for the accuracy of the various disaggregation approaches to be precisely evaluated. It is demonstrated that the correlation of transient event edges in aggregated power measurements to individual appliance transient exemplars performs better than the matching of steady-state power levels against individual appliance state combinations. Furthermore, the transient approach is shown to be the most appropriate technique for further development.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank Professor Ken Nixon for his tireless energy, sage input and consistent support at every step of my journey through this research. I would also like to thank Hugh Hunt for his encouragement and for frequently sharing the benefit of his experience without any expectation of being compensated in kind.

Funding and support from the following organisations is gratefully acknowledged:

- Eskom through the Tertiary Education Support Programme (TESP).
- The National Research Foundation (NRF).

A huge thank you is extended to my family and friends for providing me with empathetic ears throughout the course of this MSc. Most particularly my parents, Liza and Barry Penn, and Helen Walsh for their heroic patience, ceaseless encouragement and constant support.

CONTENTS

LIST OF FIGURES

NOMENCLATURE

Term	Description
NILM	Non-Intrusive Load Monitoring
TLM	Total Load Model
CPM	Complex Power Method
EEC	Event Edge Correlation
RMS	Root Mean Squared
EWA	Equal Weighting Average
UWA	Unequal Weighting Average
SS	Steady-State
LSM	Living Standards Measure

Chapter 1

OVERVIEW

1.1 Introduction

In South Africa, and many other emerging economies around the world, the balance of electrical power provision is a constant challenge. This state of affairs is only set to worsen, given both globally and locally expanding population sizes and the resulting growth in commercial, industrial and residential electricity demand. With increased public exposure to the shortfalls in electricity supply and the environmental consequences of electricity generation, residential electrical power consumers are becoming increasingly aware of the need to minimise their own consumption. Should they manage to reduce their electricity usage, residents stand to make a positive social contribution by limiting their demand on the national grid, as well as personally benefiting by reducing their electricity expenses. Demand side management of this nature offers a superior means by which to alleviate the stresses on national power grids, it being less costly to intelligently manage a load than to build new generation capacity or energy storage [1]. The NILM techniques investigated within this research are intended to contribute towards demand side management by providing tools with which consumers may analyse and modify their electricity consumption.

However, the intelligent management of power consumption is a complex undertaking. It requires that the contribution of individual appliances to the total power consumption be determined, allowing for efficient appliance usage behavioural modifications to be enacted. As many household appliances perform automated operations which are not directly controlled by the user, often becoming active at unexpected times, it is difficult to ascertain the contribution of each appliance via user observation alone. For example, a geyser or a fridge may become active at any time in order to heat or cool its contents. Detailed per appliance usage statistics allow for intelligent demand management decisions to be made, such as the installation of timer switches on geysers and fridges, the selection of economical appliance operational modes or the replacement of power hungry appliances with more moderate alternatives. However, most consumers are left with nothing more than basic intuition and their monthly electricity bill as tools with which to attempt to reduce their electricity consumption.

The desired level of per appliance electricity consumption information may be obtained by installing power monitoring devices on every appliance contained within a residence. However, this approach is financially infeasible due to the large number of monitor units required. Furthermore, many appliances are physically located such that they may be difficult to access for the purposes of monitor installation and data recovery. In contrast, NILM techniques only require measurement of the total power being delivered to the residence, which may be obtained via the installation of a single monitoring device at the point of electrical connection between the internal circuitry of the household and the outside transmission network. This is due to the fact that NILM techniques take the total power generated by all of the appliances active within the household at any time and disaggregate them in order to determine which individual appliances contributed to the total power measurements captured. Thus, NILM techniques offer a solution to the problems associated with direct per-appliance monitoring, providing the most practical and non-intrusive approach for recording the power consumption of individual appliances within a building.

Many NILM techniques are currently being developed and implemented, as reviewed in Chapter 2. Total power measurements can be complex to disaggregate, especially as increasing numbers of appliances become active within the measured circuit. None of the techniques developed thus far have overcome all of the challenges presented by complex appliance combinations and hence no single solution has emerged that is sufficiently free of error to be considered definitive. Thus it is of interest to continue research in this field, investigating both new and old NILM techniques with a view to improving their disaggregation performance.

In this particular piece of research, three NILM disaggregation techniques are comparatively evaluated. The first of the techniques, Total Load Model (TLM), is based upon the foundational NILM technique (of the same name) developed by Hart, which utilises real power steady-state measurements to make appliance identifications [2]. The second technique, Complex Power Method (CPM), utilises both real and reactive power steady-state measurements to refine the purely real power approach used in TLM. The third technique, Event Edge Correlation (EEC), utilises transient event edges in the total power measurement to identify appliances.

Real power TLM is a well researched technique, with the name 'TLM' being drawn directly from the seminal work by Hart [2]. Whilst the basic principles underlying both CPM and EEC have been previously investigated by many researchers, including Hart, their specific conceptualisation and implementation within this research are unique [2]. The names 'Complex Power Method (CPM)' and 'Event Edge Correlation (EEC)' are not known outside of this research, having been created specifically to refer to these two disaggregation techniques within this dissertation. By applying all three techniques to the same set of data, it may be ascertained which of TLM, CPM and EEC offers the best solution for disaggregating total power measurements.

1.2 Research Question

The comparative evaluation of TLM, CPM and EEC leads to the formulation of the following research question: *Does the use of steady-state level combinations, involving both real and reactive power, or the use of real power event edge correlations provide the best accuracy when disaggregating total power measurements?*

1.3 Research Area

The research is concerned with the problem of total power measurement disaggregation in a domestic household context, where appliance usage patterns may be determined by separating the total power consumption into its constituent loads. Three disaggregation techniques are comparatively evaluated such that it may be ascertained which is the most suitable for further development. The research area and scope are discussed in Sections 1.3.1 to 1.4 below.

1.3.1 Disaggregation Techniques

All three disaggregation techniques, and their specific implementations in this research, are presented in detail in Chapter 4. A brief description of TLM, CPM and EEC follows below.

TLM: This technique compares all possible combinations of individual appliance steady state real power consumption levels against the measured total real power steady-state levels in order to ascertain which appliances (or appliance states) might be active for each steady state observed within the total power measurement. This requires that the steady state real power consumption levels be known for each appliance within the circuit being monitored.

CPM: CPM is an extension of TLM that utilises reactive power steady state levels to provide an additional layer of information. As the majority of household appliances are purely resistive, most circuits will only contain a few appliances with reactive power components. By comparing all possible combinations of individual appliance reactive power steady state levels against the measured total reactive power steady state levels, several appliances may immediately be eliminated from the list of potential contributors to the total real power consumption levels. The resulting rationalised combinations of appliances are then used with the TLM method to make final identifications. This requires that both the real and reactive power steady state consumption levels be known for each appliance within the circuit.

EEC: Total power measurements contain clearly defined edges, which represent appliance operational state changes within the measured circuit. EEC compares such edges detected within the total real power measurements against the individual waveforms of each appliance within the circuit being measured. Correlation is utilised to find the closest match between the detected rising and falling edges and the corresponding samples extracted from individual appliance measurements such that the appliance state changes responsible for each edge in the measurements may be identified. The resulting series of state change events may be utilised to ascertained which appliances are active at any point within the measurement. This requires that the real power waveforms be captured for each appliance within the circuit being measured, and that leading and trailing edge samples be extracted from these waveforms.

All three techniques presented above may be implemented with standard power measurement devices, as they do not require unusually high specifications in order to be utilised. Whilst other more advanced techniques, such as frequency analysis, may offer good alternatives for investigation; they require specifications that exceed the abilities of the measurement device used in this research (see Section 3.1.2), along with the majority of measurement devices that would commonly be used in real-world household implementations.

1.3.2 Appliances Utilised in Laboratory Experiment

The appliances utilised in the laboratory experiment, presented in Chapter 3, are all relatively fundamental household appliances. This includes appliances with heating elements and DC motors, but excludes more complex appliances, such as fluorescent lights, computer power supplies or washing machines, that may either contain power shaping electronics, or exhibit more than three operational states. The appliances included in the laboratory experiments were selected such that the complexity of the measurements be reduced, allowing for effective comparative evaluation of the three disaggregation techniques. It should be noted that a real-world household would likely contain multiple instances of some of the appliances utilised in this experiment. For example, several lighting devices might be expected to be found within one household. However, only one of each appliance is utilised in this experiment. This is done to reduce complexity and to focus on comparative evaluation of the techniques, rather than attempting to directly simulate a real-world household.

1.3.3 Machine Learning

Machine learning techniques, such as Fuzzy Logic, Pattern Recognition and Artificial Neural Networks, are commonly researched in the context of the NILM field, where they are used to improve upon the performance of underlying disaggregation techniques such as TLM, as discussed in Chapter 2. Whilst this is a valid area of research, it often neglects consideration of the fundamental disaggregation approaches upon which the machine learning techniques are employed. Should a NILM technique with poor accuracy be used as the underlying method of disaggregation, its limitations will adversely affect the results returned by the machine learning technique built upon it.

This research does not involve any machine learning techniques, or other

similar logic based approaches. Rather, it is focussed upon evaluating three fundamental approaches to total power measurement disaggregation, in order to determine which of them is the most suitable for further development. The conclusions of this research could thus be of value to any researchers in the machine learning field looking to determine the best underlying NILM approach to which to apply machine learning techniques.

1.3.4 Processing Platforms

Many NILM systems intended for installation in the field require that processing be performed on embedded platforms that are physically integrated into a single unit and may be used to disaggregate total power measurements on site. Under such conditions, processing power and data storage capacity become fundamental considerations. Thus any NILM system being designed with these constraints in mind will be limited by the need to be computationally efficient. The post-processing approach followed in this research allows for the disaggregation techniques under investigation to be evaluated without consideration of this limitation.

Given the constantly accelerating evolution of embedded processing platforms, computational feasibility is gradually becoming less of a consequence for systems of this scale, opening the door for research conducted in post-processing contexts to be directly applicable to the field. Where considerations of processing power may be neglected, there exists more room for techniques and concepts to be researched in a more pure and directly scientific manner, as attempted within this research.

1.4 Expected Challenges

The disaggregation of total power measurements is not a trivial undertaking, due to the nature of the systems that produce the measurements. The

following list contains a summary of the expected challenges.

• **Similarity Between Appliances:** Many appliances, or combinations thereof, either consume similar power levels, or present similar characteristic features in power measurements. This makes it problematic to differentiate between appliance activities, where the characteristic being used to make identifications is not significantly varied between separate appliances, or appliance combinations. TLM, CPM and EEC are all subject to this consideration, where their disaggregation accuracies will be directly affected by the level of differentiation in power characteristics found between appliances.

• **Variability in Power Levels:** The power consumption levels expected for certain appliances, or appliance combinations, may vary considerably from measurement to measurement, depending on the conditions in the circuit being measured and the presence of noise. Certain appliances may also possess hidden states that are not detected in individual measurements, as their appearance is either subject to the operation of other appliances, or is affected by external factors that were not present during individual measurements. This inconsistency makes it difficult for NILM techniques using manual training schemes, as discussed in Section 2.3.1, to make accurate appliance state identifications. This includes TLM, CPM and EEC, all of which rely on manual training approaches, as presented in Chapter 4.

• **Combinatorial Approaches:** Any NILM technique based upon the generation of appliance state combinations will be subject to a high degree of difficulty when attempting to search that combination for potential matches for measured levels. This is due to the large number of combinations that may be generated from a small collection of appliances. For example, the 14 individual appliance states presented by the 8 appliances included in the laboratory experiment can produce 2048 possible combinations of appliance states, as per TLM. Taken in combination with the challenges mentioned in both of the above points, it is highly challenging to find accurate matches if some mitigation of the size of the combination is not implemented, as per CPM.

The impact of these challenges upon the accuracy of each of the three techniques is presented and discussed in Chapter 5 and Appendix G.

1.5 Research Methodology Overview

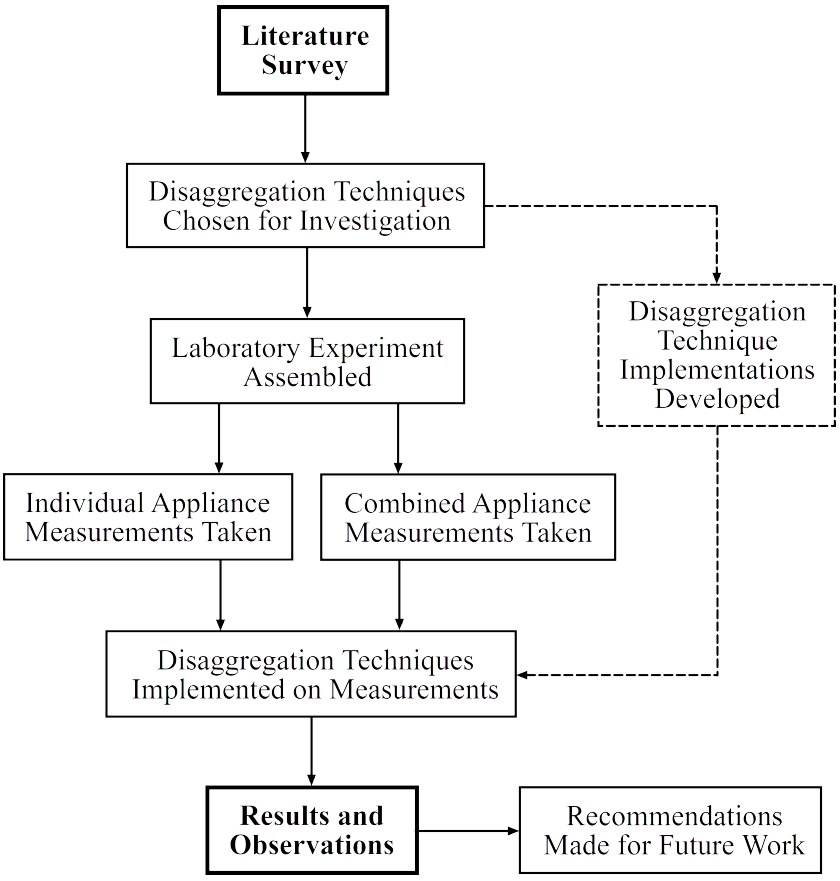

Figure 1.1: Overview of the research methodology.

An overview of the research methodology is presented in Figure 1.1. The NILM approaches and concepts most relevant to the research area covered in this dissertation are presented in Chapter 2. In order to have data with which to comparatively evaluate the three disaggregation techniques, total power measurements are required. Whilst datasets containing total power measurements from real-world households are available, these lack accompa-

nying ground truth information and hence are not appropriate for making accurate assessments of disaggregation accuracies. To obtain total power measurements and corresponding ground truth data, a circuit containing household appliances must be assembled under laboratory conditions. This allows for individual and combined appliance measurements to be performed, as discussed in Chapter 3.

Implementation of the disaggregation techniques is performed in a post-processing context, using the *MATLAB* software package [3]. The mechanics of each of the techniques is presented in full in Chapter 4, along with the *MATLAB* code written for their implementation. Underlying all three of the techniques are a number of basic measurements processing functions, created to provide the framework upon which each of the NILM implementations are built. Included amongst the processing tasks performed by these functions are the detection of appliance operational state change events and the determination of steady state power levels.

The results obtained from implementing TLM, CPM and EEC on the total power measurements are presented and discussed in Chapter 5 and Appendix G. EEC is the most promising of the techniques and the one that should be preferentially considered for further development; especially where a large number of appliances are contained within the circuit being measured. Furthermore, CPM provides a marked improvement upon the accuracy of TLM. A future research methodology for producing extensive information-rich data is presented in Chapter 6, along with a system for using previous identifications to improve EEC accuracy via the logical elimination of appliance states.

1.6 Research Justification

In the current global and local climate, where the demand for electricity often outstrips the supply, the management of residential power consumption is of

paramount importance. Residents have only very limited means by which to determine the per-appliance electricity consumption for their households. The observation of appliance usage behavioural patterns used in conjunction with electricity bills is not effective, especially given that many appliances do not conform to expected power consumption patterns; either running background operations when seemingly inactive or becoming active at unknown times. A feasible method for determining per-appliance power consumption would offer value to residents of households through utility bill savings, and to society through the efficient management of this scarce resource.

As discussed in Section 1.1, it is infeasible to place individual power monitors on every appliance within a household. NILM techniques offer a superior solution, where a potentially inexpensive device may be utilised to measure and disaggregate measurements taken from a single point of installation. The benefits that this stands to offer electricity consumers makes it worthwhile to conduct further research and development in the NILM field. Total power measurement disaggregation is difficult to achieve with high accuracy, especially in cases where a large number of appliances are contained within the circuits being measured. However, NILM systems have to provide accurate results in order to be of value, as electricity consumers are unlikely to adopt such systems if the appliance identifications that they produce are questionable. Due to the considerable challenges that must be overcome in order to accurately disaggregate total power measurements, the NILM field remains an area of ongoing research.

1.7 Organisation of Dissertation

This dissertation contains a further six chapters and seven appendices. The content of these chapters and appendices are presented in summarised form below.

Chapter 2: A review of the NILM field is presented, with an emphasis on

steady state and transient signature analysis as relevant to the three disaggregation techniques being comparatively evaluated within this research. Alternative NILM approaches are also briefly discussed.

Chapter 3: The entire measurement process is presented. This includes discussion of the appliances that constitute the electrical circuit assembled for the experiment, the individual and combined appliance measurements captured and the measuring device utilised.

Chapter 4: The conceptual mechanics and software implementations of TLM, CPM and EEC are presented, along with the underlying measurements processing functions that provide the basic platforms required for the implementation of the disaggregation techniques.

Chapter 5: The results of applying the three disaggregation techniques to the measurements are presented and discussed. This includes disclosure of the scoring methods used for each set of results and their effect on the comparative evaluations. A set of observations pertaining to the performance of each technique is condensed out of the results analysis.

Chapter 6: Recommendations for future work in the same research area are made. This includes the proposal of an improved methodology for the measurement process, and the use of previous identifications to refine accuracy.

Chapter 7: The conclusion of the dissertation is presented. This includes a brief overview of the research process, a summary of the results of the investigation and the observations that they gave rise to, along with recommendations for future research.

Appendix A: Real and reactive power plots that support both the individual and combined appliance measurements are presented.

Appendix B: The ground truth data that accompanies the combined measurements is presented.

Appendix C: A selection of the *MATLAB* functions developed for the implementation of the disaggregation techniques are presented.

Appendix D: The complete set of results produced by applying TLM to the combined measurements are presented.

Appendix E: The complete set of results produced by applying CPM to the combined measurements are presented.

Appendix F: The complete set of results produced by applying EEC to the combined measurements are presented.

Appendix G: Detailed analysis and discussion of the results contained in Appendices D through F is conducted.

Chapter 2

BACKGROUND

A brief overview of NILM is conducted, followed by an introduction to appliance signatures and their components. Basic steady-state signature disaggregation techniques are discussed, including TLM and the concepts of manual and automatic training. Challenges facing TLM and other steady-state techniques are presented. Transient signature disaggregation is discussed, including the features which may constitute transient waveforms and the NILM technique of direct transient feature classification. This is followed by brief discussions of alternative NILM techniques and disaggregation datasets.

2.1 NILM Overview

A single measurement point, usually located where electricity enters a household from the outside grid, may be used in order to perform an aggregated measurement of all power being consumed within a residence. NILM involves the disaggregation of this total power measurement such that the individual

operations of appliances contained within the household may be identified. This approach to per appliance power consumption monitoring does not require any additional sensors or measurement equipment beyond the single device used to capture the total power measurements. In comparison to alternative approaches, such as the direct observation of the appliance usage behaviour of residents and the installation of separate power monitors on all appliances, NILM is both less intrusive in nature and more physically feasible to implement [4].

The cost effectiveness of NILM systems, and their single point of installation, allows for researchers to access a larger sample of households for the same cost and labour than other methods. The non-intrusive nature of NILM helps to minimise any observation biases that might otherwise affect the power consumption behaviour of residents taking part in research experiments. Thus the data provided by NILM systems has the potential to be more relevant to real-world households than that gathered using other experimental approaches, and may cover a larger sample of households across a wider range of socio-economic strata.

There are three main processes that must be enacted within any NILM system [5]:

• **Measurement:** Power measurements must be performed at the chosen point of installation.

• **Event Detection:** Important events, features and characteristics must be detected and extracted from the power measurement.

• **Identification:** The known qualities of previously measured appliances must be utilised in order to identify these events, features and characteristics such that they may be associated with a particular appliance.

Whilst measurement and event detection are fundamental components of any NILM technique, the main differences between NILM techniques are

found in the identification phase; where alternative approaches may be used to match any events, features or characteristics detected within the power measurement to the operations of individual appliances.

2.2 Appliance Signatures

The various events, features and characteristics that may be found in total power measurements make up appliance 'signatures' that may be attributed to the operation of each of the appliances contained within a household. For example, variations in power consumption (real and reactive) and current harmonic characteristics are closely linked to the nature of the load, and this information may be used to identify loads from out of the aggregated data [6]. There are two broad categories of signatures that may be considered, steady-state and transient [7]. Steady-state signatures are constantly present in the measurements whilst appliances are operational, whereas transient signatures only appear for short durations; such as the transitions between appliance operational states (e.g: when switching appliances between 'on' and 'off' states) [8]. When a signature has been extracted from a power measurement it may be broken up into the following components, all of which can be used for identification purposes [9], [10]:

• **Levels:** The steady-state power levels that exist between appliance operational state changes where no other features or events are found in the power measurement.

• **Edges:** The edges of the signature, which are seen when the appliance changes operational states. These may be step changes, or could come in a variety of different shapes, depending on the internal electronics and the physical operations of the appliance in question.

• **Sequences:** The sequence of power level changes observed between edges. This indicates that an appliance is passing between a number of different

states (e.g: on a washing machine; the wash, drain and spin cycles). With some appliances these may vary with different iterations of usage. For example, residents may have to choose between a variety of settings, or automatic controls might respond to some form of feedback, such as temperature.

• **Trends:** Changes in power levels that are not edges between operational state changes. Expected shapes include transient peaks, pulses, oscillations, vibrations and slopes. They are distinct from edges in that they either don't clearly connect two steady-state power levels, or are of a significant duration to represent more than a transition in operational state.

• **Time:** The time and date, as well as the duration, of the appearance of an appliance signature in the power measurement. Many appliances have fixed operational periods, which can aid in their identification. Furthermore, the specific times at which appliances become operational can also offer clues to assist in identifying their contribution to the total power measurement (e.g: a large power consumption increase seen during cooking times could be attributed to the activation of an oven).

• **Electrical Circuit:** Knowledge of which electrical circuit the particular measurement comes from can assist in the identification process, provided that it is known where in the household each appliance is installed. Unfortunately this requires the installation of further sensors, one for each circuit on the distribution board. The gathering of information on the installation locations of individual appliances and the use of multiple measurement points is a violation of the underlying principles of NILM. Thus this factor would not be included for consideration in a pure NILM system. Furthermore, many appliances are regularly moved around to different locations within the household (e.g: hair-dryer), making the use of a predefined set of circuit-appliance assignments problematic.

A variety of different techniques may be employed in order to extract these components from total power measurements and make appliance identifications. Such techniques utilise either the steady-state or transient power char-

acteristics of the appliances contained within the circuit being measured, the basic principles of which are discussed in sections 2.3 and 2.4 respectively.

2.3 Steady-State Signature Disaggregation

The steady-state signature of an individual appliance is the pattern of its steady-state property of interest (e.g: power consumption level) during appliance operation, be it for purely 'on/off', or multiple state appliances [11]. This includes periods of consistent operation and transitions between states, the latter being identified via the presence of edges or ramps within the total power measurement. The more distinct the transitions are, separating the measured data into a series of easily distinguishable step change events, the more reliably it may be determined which particular loads are present in the total measurement at any point in time [12]. Steady-state events are far simpler to capture than transients, as they are present in power measurements for longer periods of time. Furthermore, they are intuitively additive, making two overlapping signatures far easier to disaggregate than simultaneously occurring transient features [2].

However, the use of steady-state signatures does not come without problems. They can be difficult to distinguish in cases where they either overlap ambiguously or change state in rapid succession, making disaggregation of the total measurement virtually impossible [13]. Consequently, their use may need to be supplemented with other techniques if an accurate disaggregation solution is to be obtained.

2.3.1 Total Load Model (TLM)

The version of TLM discussed in this chapter is the most basic, concentrating only on real power levels. Hart did develop TLM further, to include both real and reactive power, as discussed in his 1989 and 1992 papers [8], [2]. However,

the exclusive consideration of real power allows for this cornerstone of steady-state approaches to be explored at a fundamental level. The technique is primarily focussed on step changes in the total real power measurement, which may be used to determine the combination of appliance operational states responsible for each measured total power level. In order to simplify the process, Hart originally considered only two states per appliance, limiting all appliances to 'on' and 'off' modes and avoiding multi-state appliances.

Within any household being measured, each appliance will be wired in parallel to the incoming power bus as shown in Figure 2.1.

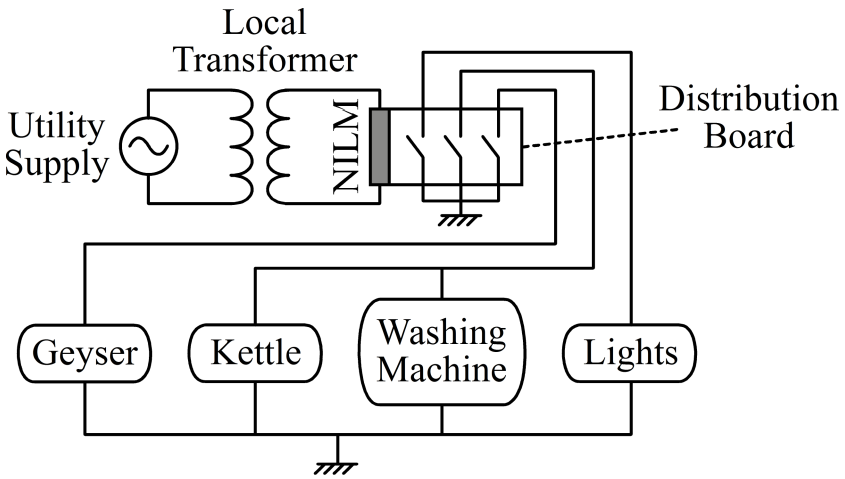

Figure 2.1: Sample household circuit layout with NILM system installation.

Due to this parallel configuration, the total power measurement is constituted by the sum of the power consumed by all of the appliances connected within the household. The total load model can be mathematically expressed as shown in Equation 2.1 [2].

$$P(t) = \sum_{i=1}^{n} a_i(t) P_i + e(t) \tag{2.1}$$

Where:

$P(t)$ is the total power measured at any time t.
$a_i(t)$ is an n-component Boolean vector containing the 'on/off' state (represented by 0 or 1) of each appliance at any time t.
P_i is the vector of the power consumed by each appliance when in the 'on' state.
$e(t)$ is the error or noise present in the system at any time t.

From this model the operational state of every individual appliance in the household may be estimated, given that its power consumption is known. In order to ascertain this, the appliance state combination vector needs to be populated. This vector is the combination of operational states that results in $|e(t)|$ being at a minimum, which can be expressed as shown in Equation 2.2 [2].

$$\hat{a}(t) = argmin_a |P(t) - \sum_{i=1}^{n} a_i P_i| \tag{2.2}$$

Where:

$\hat{a}(t)$ is the optimised, or 'best-fit', 'on/off' appliance state vector at any time t.
$P(t)$ is the total power measured at any time t.
a_i is an n-component Boolean vector containing the 'on/off' state (represented by 0 or 1) of each appliance.
P_i is the vector of the power consumed by each appliance when in the 'on' state.
n is the number of appliances included in the system.

To find where appliance states have changed, significant positive and negative step changes within the total power measurement must be detected [8]. When each new steady-state level is reached, the optimal state vector must be found that satisfies Equation 2.2. This process allows for every step change in the total power measurement to be assigned to a state switching event for one of the appliances within the household, producing a time-line of appliance state

transitions. In order to find the optimal vector after each switching event, the power consumption levels of each appliance state must be known. It is not sufficient to rely upon rated power values, as these are given only in terms of real power and are often approximations. Thus new measurements must be made in the context of the installed NILM system, whereby the effect of each appliance state on the total power measurement may be recorded. This process is known as 'training', and may be performed via manual or automatic methods.

Manual Training

Manual training requires that all of the appliances within the circuit being measured are individually activated and deactivated, so that the effects of the state transitions of each appliance on the measured total power measurement may be observed and recorded. Due to error and noise ($e(t)$), the step changes observed when repeatedly switching the same appliance 'on' may not be the same for each iteration. Thus statistical clustering analysis techniques may be applied to the results in order to identify distinct regions on the complex power plane that conform to each of the appliances connected to the circuit. From these regions, a lookup table can be generated that allows for identification of appliance state changes during normal operation of the measurement system [8]. Whilst this process violates the NILM philosophy by virtue of being fundamentally intrusive, it is a once-off operation that could be considered to be part of the installation process. Once completed, the system can begin performing its function without further intrusions being required.

Automatic Training

Under the automatic training scheme proposed by Hart, the system is installed and measurements are taken without performing any sort of preparatory process [2]. Once data has been collected over a sufficient period of

time for all appliances within the household to have been operated, statistical methods are used to identify steady-states. These steady-states must then be assigned to a particular appliance by determining which load class or consumption level best fits its power characteristics. Where complex power is considered, certain appliances include power factor correcting elements (e.g: fluorescent bulb) that are incorporated in order to reduce their reactive power. Since this makes them appear purely resistive, it could lead to them being classified incorrectly [14]. The time at which each steady-state is measured may also be factored into the identification process. For example, a kettle would likely not be utilised outside of waking hours, yet a geyser might be active intermittently over a 24 hour period. Similarly, duty cycle data may be used to aid in the discrimination process, as many appliances switch 'on' and 'off' periodically in order to control some factor, such as temperature [8]. Using such information, of which the power consumption level is paramount, the initial automatic training set of data may be employed to build a lookup table for appliance identifications during normal operation.

Additional information notwithstanding, the automatic training approach is still likely to lead to considerable levels of identification inaccuracy where appliances with similar reactance, power magnitude and timing characteristics are concerned. This expected error must be balanced against the non-intrusive nature of the method when deciding whether to adopt it ahead of the manual alternative. Certainly, the automatic approach embodies the philosophy of NILM far more closely and this alone may be reason enough to pursue it preferentially. The automatic training approach described in the 1992 paper by Hart is only partially automatic, as it requires a manual analysis of the initial data in order to populate the lookup table [2]. However, the potential does exist to automate this process via the use of predefined classes, which would render it fully automatic in nature. The remaining problem with this training approach is that residents may not utilise all of their appliances within the initial period. This would mean that those appliances would not be included when attempting to optimise the appliance state vector, introducing further error into the results.

2.3.2 TLM Shortcomings

TLM features a number of shortcomings that impede its ability to provide an effective solution for real-world applications, as listed below. These same shortcomings also affect other steady-state techniques that operate upon the same underlying principles as TLM, including CPM.

- Similar Power Consumption Levels
- Reliance on Steady-States
- Optimisation Processing
- Addition of Appliances
- 'Best Fit' Approach
- Multiple Appliance States
- Appliance States Transitions
- Continuously Variable Appliances
- Simultaneous Events
- Negative Real Power
- Non-Linearity

Similar Power Consumption Levels

Different loads may exhibit almost identical power consumption levels, regardless of their load class and function. This makes them difficult to distinguish from one another, especially in the context of noise within the circuit being measured. Furthermore, an appliance may consume a certain amount of power under one set of conditions in the circuit, but a slightly different amount under another, contributing to the difficulty. As the number of appliances with similar power consumption levels increases, so the task of distinguishing between them becomes more challenging.

Reliance on Steady-States

TLM relies on the existence of steady-state periods between state switching events, where appliance behaviour is constant and no other appliances undergo transitions. However, it is not a given that such periods will be found between every appliance state change event. This is especially so when a large number of appliances are contained within the household, bearing in mind that their power consumption behaviour may include ramps, ripples, oscillations and other features that reduce the likelihood of experiencing steady-state periods. Furthermore, some appliances can take long periods to make transitions, in the order of seconds or minutes (e.g: a large fan speeding up to final velocity), whereas others may never reach a steady-state at all (e.g: a variable speed drive).

Optimisation Processing

The combinatorial optimisation problem expressed by the total load model may only be solved if n is small, otherwise the amount of processing required to test all combinations is impractical. Hart recommends the use of heuristic algorithms for this application [2]. However, whilst this approach may be practical in the context of a manual training scenario where the exact number of appliances is known, it is more problematic under the automatic training paradigm with unknown n. In addition, should multiple appliance states (more than binary) be considered, the processing requirements are further increased.

Addition of Appliances

Once the training approach is completed, be it either manual or automatic, the operation of any new appliances that are added into the household will be interpreted by the optimisation algorithm to be a combination of other known

appliances [2]. Given that this is a feasible event in real-world households, the inability of the system to 'learn' any further appliances without manual intervention is problematic and could introduce unacceptable levels of error into the results.

'Best Fit' Approach

A further issue concerns the optimisation technique used to find the 'best fit' appliance state vector. In many cases this method can lead to unrealistic results, as illustrated in the following example taken from the 1992 paper by Hart [2]:

In a household that contains four appliances, at time t the power (considering only real power in this example) for each appliance is:

$P_1 = 100\text{W}$ $\qquad$ $P_3 = 300\text{W}$
$P_2 = 200\text{W}$ $\qquad$ $P_4 = 401\text{W}$

For a total power measurement value of 500W, this gives a best fit of: $\hat{a}(t) = [0, 1, 1, 0]$

However, if the power has changed at time $t + \Delta t$ to 501W, then the best fit becomes: $\hat{a}(t + \Delta t) = [1, 0, 0, 1]$

This implies that there has been a simultaneous change in the operational state of all the appliances, which is improbable. Rather, the change is more likely to be expressed by some noise, in this case: $e(t + \Delta t) = 1$

In order to minimise the number of erroneous assignments made in this fashion, thresholds need to be set such that the most probable appliance combination is not recalculated unless the total power change exceeds the maximum expected noise level. This approach compensates for noise related errors without resulting in incorrect appliance state determinations, provided that the noise remains within the predefined thresholds. However, the use of

thresholds may lead to the exclusion of appliances that only consume very low levels of power (e.g: a cellphone charger) from the list of appliances that the system can identify.

Multiple Appliance States

Many appliances have multiple operational states, not just 'on' and 'off' (e.g: a washing machine), for which this disaggregation method does not cater. Each of the multiple operational modes is characterised by a particular level of power consumption, thus there are many more potential appliance combinations that could accompany every steady-state total power level than catered for by the binary ('on/off') approach. If all of the possible appliance states are included, the resulting number of potential combinations may make the $a(t)$ vector so long that it may not be feasible to process it. Furthermore, the probability of erroneous conclusions being reached by the algorithm when considering so many possibilities becomes too high for the system to be effective.

Appliance State Transitions

Not only do many appliances have multiple operational states, but transitional stages may also be found to occur between steady-state conditions. For example, Figure 2.2 shows the power consumption for a purely resistive toaster that switches between three heat settings; warm (low heat), crispy (medium heat) and burnt (high heat).

During such transitional stages, new appliance state vector combinations will be calculated every time the power level exceeds the threshold set for the identification of appliance state transitions. However, the appliance in question is still undergoing a transition at this point, not yet having reached the new steady-state. This results in a series of erroneous appliance state combinations being assigned to that transitional section of the measurement.

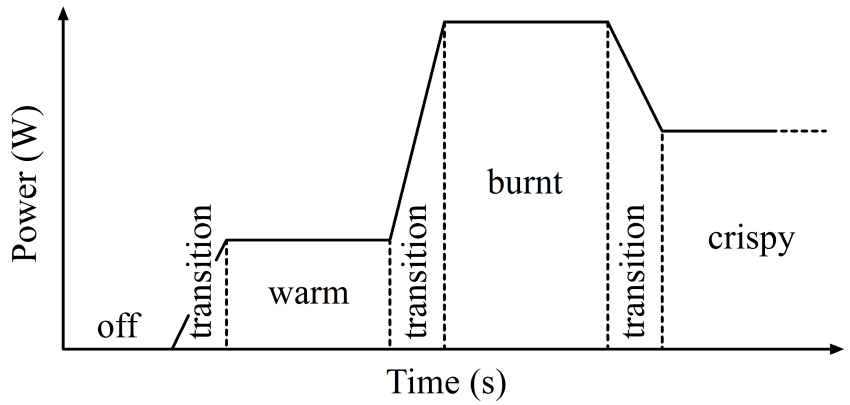

Figure 2.2: Steady-state power consumption and state transitions for a toaster with three settings.

Continuously Variable Appliances

Continuously variable appliances never settle into distinct states, except when 'off', and thus cannot be catered for by the model. As the power consumption levels for such appliances never reach a steady-state, this causes the same problems experienced with state transitions. Whenever a continuously variable appliance is introduced into a circuit, its operation increases the error levels considerably. Given that many households contain these sorts of appliances (e.g: a power drill), the inability of the model to compensate for their use is problematic.

Simultaneous Events

If two or more appliances are activated simultaneously, it could appear as if a single appliance with a larger power consumption level has become operational. The system may not be able to discern between the two loads and could instead ascribe the event to a single appliance possessing the sum

of the power consumption of the actual appliances involved [14]. In order to minimise this error, there should be a significant steady-state power period between appliance state transitions that allows for individual appliance activation events to be distinguished. The minimum length of such a period may be determined using statistical algorithms, filtering, differentiating or peak detection [5]. The sampling rate used to perform the total power measurement will influence the probability of simultaneously events being found, where the higher the sampling rate, the lower the probability of events appearing to be simultaneous. As sampling rates approach infinitely high speeds, so the probability of encountering simultaneous events tends towards zero.

Negative Real Power

An intrinsic assumption of the total load model is that the operating real power consumption of every appliance is never negative, i.e: that no appliances generate power [2]. Whilst this may seem to be a reasonable assumption to adopt for households, it does not hold in cases where renewable energy sources are included in the system. Recent years have seen increased interest in renewable energy and microgrid implementations within households, posing a further challenge to the real-world applicability of TLM.

Non-Linearity

TLM is theoretically only applicable to linear appliances. Whilst it may be generalised for non-linear appliances, as shown in Equation 2.3, Hart found that the linear version of the model (with $\beta = 2$) still produced the best results in field trials [2].

$$P_{norm}(t) = \left(\frac{230}{V(t)}\right)^{\beta} P(t) \tag{2.3}$$

Where:

β is the order of linearity (e.g: $\beta = 2$ is linear).
$V(t)$ is the RMS voltage measured at any time t.
$P(t)$ is the total power measured at any time t.
$P_{norm}(t)$ is the normalised power at any time t.

If separate exponents are included for the real and reactive components of the load, the normalised power signatures can be statistically clustered such that each of the resulting clusters may be identified based upon its real and reactive power consumption, as discussed in Section 2.3.1 [2].

This concludes the discussion of TLM shortcomings, all which affect steady-state techniques based upon similar principles as TLM, including CPM. None of the appliances incorporated into the laboratory experiment generate power, thus the negative power (Section 2.3.2) shortcoming is not applicable to this research. Furthermore, the post-processing approach utilised for the implementation of TLM and CPM also makes the optimisation processing issue (Section 2.3.2) non applicable, as discussed in Section 1.3.4. However, these two shortcomings remain relevant to the discussion of TLM, CPM, and other similar steady-state techniques, as they may be present in other NILM implementations.

2.3.3 Steady-State Event Detection

Rapid changes in the variables of interest (e.g: power consumption) produce defined edges in the measurement, which may be taken to denote appliance state transitions. Before each edge may be classified as representative of a state transition event, it must be ascertained whether it is truly a waveform edge, rather than a peak in the noise. In order to perform statistically verified event detection, the distribution of the most recent measurements must be compared to that of those captured in the previous iteration of sampling [15].

Alternatively, event detection thresholds may be used in order to disregard

any steady-state level change events that fall below selected minimum values. This prevents any fluctuations occurring beneath these levels from being treated as events, minimising the impact of noise on the effectiveness of the system. Threshold values may be chosen either nominally (based on system expectations), or during the course of the training process. This threshold based approach to event detection does pose a problem with regard to low power consumption appliances (e.g: a cellphone charger), which may fall below the threshold values and thus be excluded from the possibility of identification.

The Generalised Likelihood Ratio (GLR) is sometimes used to identify edges in steady-state time-series data, especially under conditions of noise and other distortions [16]. GLR algorithms calculate a decision statistic that utilises probability distributions applied to the data recorded both before and after a potential change in mean (i.e: across a window), in order to rule whether or not a state transition actually occurred [16].

This concludes the discussion of steady-state signature disaggregation techniques as they apply to the research presented in this dissertation. A similar discussion is conducted for transient signatures techniques in Section 2.4 below.

2.4 Transient Signature Disaggregation

The transient behaviour of loads may be defined as the effect that their operation has on the electrical waveforms that pass through them. Peculiarities of the design and operations of appliances, as well as the electrical components that they contain, introduce distinguishing features into the waveforms which may be used to identify the operation of individual appliances [17]. For example, the transients produced when a laptop computer is activated will be significantly shaped by the charging of the capacitors in its power supply, whereas a heating element with no added electronic components will

introduce a different set of features into power measurements.

Transients may vary between measurements, and are affected by the point in the voltage cycle where appliance state changes occur, making them challenging to analyse [2]. Switches, both mechanical and electromechanical, are a common source of transients due to the physical manner in which their contacts interact during opening and closing operations. Wherever switches are used, they have the potential to introduce bouncing, rocking, sliding or other deformations into the power measurement [18]. Given that most appliance operational state changes will be accompanied by switching of some nature, the detection of transients can provide valuable information to aide in the identification of individual appliances and their various state transitions.

Whilst transient signatures are more difficult to detect than steady-state alternatives, and may provide information that is less directly relevant to per appliance power consumption, they can be very useful when used in conjunction with steady-state signature techniques [11]. Furthermore, some appliances continuously produce transient features during the course of their normal operation, by which they may be identified (e.g: the commutator of a motor) [18].

2.4.1 Transient Features

Transients may be characterised according to their size, duration, time constants and parametric variables. Four categories of transients that may be observed in power measurements are identified by Hart in his 1992 paper, as listed below [2].

- Ripple, Ramp and Edge
- Short Variable
- Starting Plateau
- Starting Peak

Ripple, Ramp and Edge

The current waveforms produced when appliances are activated may contain a combination of ramps, ripples and other features that precede steady-state operation [12]. These features can be processed into regions of ramps (slopes), ripples and edges, such as those shown in Figure 2.3. This allows for steady-state characteristics to be ignored and the series of transient features used to identify the unknown appliance by comparing them against a library of exemplars.

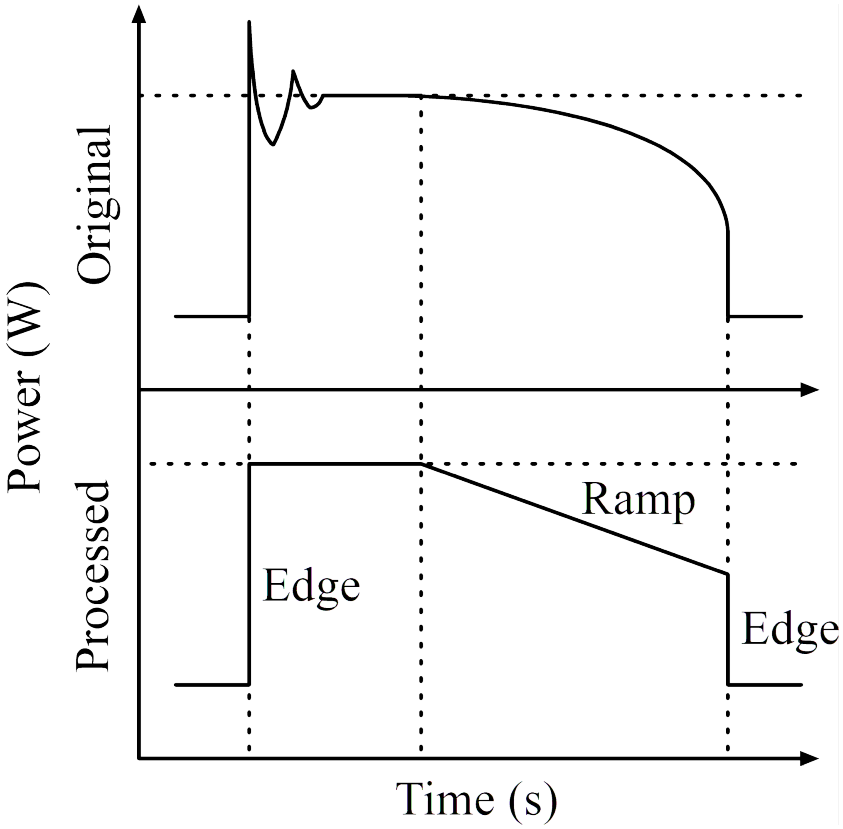

Figure 2.3: Original and processed power waveform showing edge and ramp features.

The peak of the leading edge of the appliance waveform is usually higher than the steady-state level that follows. Thus, if edge, ripple and ramp transients are extracted from power measurements, they may be processed

such that they match up intuitively, producing a series of features that return the system to the original level of power consumption, as shown in Figure 2.3 [19]. Different load classes exhibit particular operational trends which are reflected in the features constituting the processed appliance waveform. Where these trends are identified, they may be used to classify the appliances that produced them. In many cases the specific identity of the responsible appliance may be ascertained.

For example, consider an appliance that winds a spring slowly. As the tension increases, so does the current drawn by the motor. Thus a gradual ramp in power consumption is seen over the winding period, providing insight as to the nature of the appliance. Hart uses a washing machine to illustrate how transient feature sequences can offer clues to the identity of the appliances that produce them [2]:

- The appliance creates a ripple during the agitation process, due to the reversed movements of the drum required for the washing action.

- Ramps and edges would be witnessed as the various motors within the machine are activated, such as the drum and pump motors.

- An ascending ramp may occur during filling cycles, where a head of water accumulates above the pump, requiring more pressure as the process continues.

- The sequence of these features should be consistent with the cycles followed by the machine, such as the wash, drain and spin cycles.

Short Variable

Short variable transients, with periods that are complete within one or two cycles of the voltage, are found to occur across the appliance classes. They include surges and decays in current, which accompany state switching events, and can be found to occur in any number of appliances due to the particulars

of their internal circuitry. An example would be the short peaks, or surges, that result from the heating up of incandescent light bulb filaments when initially switched on. Many appliances, regardless of their load class, contain reactive electronic components that produce distinctive short variable transients when switching events occur. These can provide useful signatures for identifying appliance activities in the total power measurements.

Starting Plateau

Appliances containing motors often posses a coil which provides starting torque, but is then switched 'off' once the motor is under way. This produces a transient with a flat character, a plateau, that corresponds to the power consumed by the starting coil. This power level rapidly steps down to the steady-state power level once the motor has entered normal operation.

Starting Peak

Other motor variants draw increasing levels of power, either in discrete increments or smoother curves, as the shaft overcomes inertia and begins rotation. Once the desired speed has been reached, power consumption decays exponentially as the motor enters normal operation. This behaviour is reflected in power measurements as a smooth or stepped peak that indicates that such an appliance has initiated operation or changed state, requiring additional torque.

This concludes the discussion of transient features. Such features have particular bearing on EEC, and any disaggregation techniques which use transient events to identify appliance operations within total power measurements. Once transient features have been extracted from the measurements, they may be identified via comparison to libraries of individual appliance exemplars, as discussed in Section 2.4.2 below.

2.4.2 Direct Transient Feature Comparison

Transient features may be identified by means of direct comparison, where their curves are matched against a library of exemplars on a point-by-point basis. When selecting sampling rates for use with direct transient feature comparison, the primary consideration is the level of detail to be reproduced in the measurements; where higher sampling rates will capture greater levels of detail. However, higher sampling rates have greater data processing and storage requirements, and thus a balance of sampling rates and data specifications must be found that is most appropriate for each NILM system being developed. For direct transient feature comparison to produce accurate results, the individual appliances in the circuit being disaggregated must have distinctive transient patterns, so that they may be distinguishable from one another. A large and complex database of exemplars is required, given that transient features can be quite varied and thus need a substantial library in order to be accurately matched [13]. However, this does offer the advantage of allowing individual appliances to be more precisely identified than may be done under many other classification schemes, where signatures are merely assigned to general load classes.

There are three main methods by which transients are commonly classified or identified. In the first approach, every point on the measured transient curve is directly matched against a pre-compiled database until the 'best fit' exemplar waveform is found. The second method is more sophisticated, involving the comparison of each measured transient to the most complicated exemplars before proceeding to the simpler alternatives. This is done to ensure that a series of small sample waveforms from the library are not matched to separate sections of a measured transient, when the whole curve might be better matched to a single exemplar. By proceeding in this manner, a single transient event may be prevented from being erroneously classified as a series of distinct appliance state transitions [6], [7]. The third approach is to evaluate all of the exemplars in the database and short-list any that match sections of the measured transient. Once this process has been completed,

various combinations of the short-listed library samples can be evaluated against one another for best-fit to the measured feature [6], [7]. This ensures that sufficient classification possibilities are considered, rather than having the algorithm settle on the first acceptable match found. Whilst the second and third approaches are more exhaustive than the first approach, they have the disadvantage of requiring greater computational resources, which makes them less feasible to implement in real-world NILM systems.

The actual timing of transient events can prove problematic for classification techniques, especially where significant feature overlaps occur, as coincident transient waveforms are difficult to distinguish [6]. In their 2003 paper, Laughman et al present a method for separating overlapping transients; each library exemplar is cropped into multiple sections which are individually matched to the incoming data, where time shifts, offsets and gains are calculated using the least squares criterion [17]. This is a sensible approach, given that the overlapping waveforms are an accumulation of individual curve sections. However, the number of possible combinations that must be considered in order to determine the composition of each waveform introduces error into the classification results. In addition, such routines are computationally expensive, and thus problematic for real-world NILM system implementations.

2.4.3 Transient Event Detection

In comparison to the detection of steady-states, transients are easy to detect given that they are events or features in power measurements. Edges are the primary transient of interest in terms of event detection, as they are commonly found to accompany appliance state transitions. Edges may be easily detected, given that they are rapid changes in power consumption that appear as near vertical features in the total power measurement, followed either by a new steady-state power level or other transient feature.

Ramps are more complex to detect, as they must be approximated from the

conclusion of an edge until a stable steady-state power region is reached. Ramps may have extended durations (e.g: a gradually accelerating motor), making their detection challenging in cases where other state transitions or transient features occur within the same period, introducing steps or gradient changes into the slope. Transient event detection algorithms need to be designed such that they do not misidentify ramps as a series of other features.

Where edge detection is employed to supplement transient waveform extraction, it is possible to distinguish appliance state-change transients from other features in the power measurement. This can prove advantageous, given that the power transients that accompany appliance state transitions represent the expected nature of the appliances that produced them, and thus are more easily matched or analysed [20]. Such an approach also reduces both the required size of the exemplar library and the computational burden posed by the classification process.

2.5 Other NILM Approaches

Many disaggregation techniques have been used within the NILM field. A selection of techniques that have not been covered earlier in this chapter are briefly presented below. Pure classification processing approaches, such as Statistical Clustering, Artificial Neural Networks, Nearest Neighbour and Fuzzy Logic techniques are not included in this list.

Higher Harmonics: Higher harmonics are generated by distortions, non-linearities or power electronics contained within loads. The presence of higher harmonics in the total power measurement can provide insight into which appliances are operational, as the distribution of harmonics may be associated with particular load classes or specific appliances [17]. Higher harmonics are found by decomposing power measurements into their frequency components using techniques such as Fourier and Wavelets Transforms.

Feature Recognition: Various current waveform factors can be calculated from transients, providing a set of parameters that allow the total power measurement to be compared to individual appliance exemplars. The factors of interest may include peak, average and RMS current, crest and form factors and peak to average ratios [21].

Instantaneous Admittance Waveforms (IAW): Appliances are typically connected into household circuits in parallel, making their admittances additive in terms of the total measurement. Individual appliance IAW may be differentiated by the oscillations, peaks and distortions that they contain, allowing for identification of the appliances that generate them [22].

Phase Shift: The degree of inductance or capacitance contained within a load is expressed by the phase shift that is found in the power measurement, where positive and negative shifts indicate inductive and capacitive reactances respectively [23]. This information may be used to identify appliances based on prior knowledge of their characteristics.

Eigenvalue Analysis: If sections of the measured total power data are arranged in matrix form, eigenvalues may be found via singular value decomposition [24]. Every edge or steady-state in the power measurement, depending on the disaggregation approach being applied, may thus produce an eigenvalue that can be used to associate it with a particular appliance based on values captured during manual training.

Pattern Recognition: The shape of the variations in steady-state power levels resulting from the operation of individual appliances produce patterns within the total power measurement [11], [25]. Pattern recognition techniques may be used to extract these patterns from the total power measurement, indicating the presence of whichever appliance is associated with the pattern detected.

VI Trajectories: Symmetrical 'trajectories' are produced by plotting instantaneous current and voltage measurements captured during the operation of an appliance. Various parameters of the shape of each trajectory

contain information that may be used to describe the nature of the load being measured, providing a means by which appliance identifications may be made [26].

For each of the techniques presented above, the electrical characteristic of interest is affected by the simultaneous operation of appliances. This results in the same combinatorial problems faced by TLM and CPM, and thus the additional techniques offer no clear advantage when used alone. However, if any of these approaches were to be implemented in combination with TPM and CPM, benefit might be gained via the addition of layers of information which could supplement the combinatorial optimisation process.

2.6 Basic Load Classes

A number of basic load classes may be defined for household appliances, based on their power consumption behaviour, and may be useful for aiding in the identification of appliance operations within total power measurements [27], [10].

Resistive: Appliances that are purely resistive, e.g: kettle. These are characterised by no switch-on transient (or a very short transient, smaller than the 50Hz period of the current signal) and no harmonic content contained within the current.

Pump: Appliances containing electric motors that drive a pump, e.g: washing machine. Such devices are characterised by substantial reactive power, long switch-on transients and odd-numbered harmonic currents.

Motor-Driven: Appliances containing electric motors that are not being used to drive pumps, e.g: drill. These feature smaller turn-on transients than the pump class.

Electronically Fed: Appliances that operate at a low level of power consumption, e.g: television. Such devices are characterised by short, high-amplitude switch-on transients and contain many harmonic components within the current.

Electronic Power Control: Appliances whose operation is electronically controlled in order to operate as desired, e.g: halogen lights. Their characteristics often vary with the power level at which the appliance operates, making them difficult to consolidate within a single class.

Fluorescent Lights: Fluorescent light bulbs. These are characterised by a long two-step switch-on transient, a very high amplitude third harmonic current and a significant phase shift between the voltage and current.

Another set of categories based on a slightly different set of behavioural aspects may also be considered [27]:

Permanent: Appliances that are constantly on, e.g: alarm systems.

On/Off: Appliances that feature only two power consumption states, namely 'on' and 'off' with minimal features or events occurring between the state transitions, e.g: toaster (similar to resistive class above).

Finite State: Appliances that pass through more than a single state of power consumption during operation, e.g: washing machine (wash, drain and spin).

Continuously Variable: Appliances that feature variable power consumption that does not change in discrete steps, e.g: drill.

These categories are useful for performing the classification of measured signals, where each event or feature within the data may be assigned to one of the above classes based upon its characteristics.

2.7 Disaggregation Datasets

A number of household electricity consumption datasets exist that may be used in preference to the gathering of unique data via an experimental procedure. A few key datasets pertaining to the research area covered in this dissertation are presented below, along with discussion of their drawbacks in the context of this research, providing justification for the measurement process discussed in Chapter 3.

BLUED: The Building-Level fUlly-labelled dataset for Electricity Disaggregation (BLUED) consists of aggregated voltage and current measurements from a single domestic household, sampled at 12kHz over the period of a week [28]. The operational state transitions of each appliance are recorded individually via the use of plug-level meters, environmental sensors and circuit panel meters. The inclusion of this ground truth data should make the BLUED dataset valuable for the effective evaluation of total power disaggregation techniques. However, the ground truth data was only collected at an estimated 95% level of accuracy, due to incorrect circuit tracing, appliance relocations and appliance additions during the experiment. Furthermore, approximately 25% of the appliances in the household did not register events due to low power consumption or short operational durations [28]. These inaccuracies limit the value of the dataset significantly, especially in the context of the research presented in this dissertation, where absolute ground truth data is required for the accurate assessment of all appliance identifications performed.

REDD: The Reference Energy Disaggregation Dataset (REDD) contains high frequency total current and voltage measurements from six households sampled at 15kHz over a period of several months, with an inventory of appliances included for each residence [29]. Whilst this is a fantastic resource for testing disaggregation techniques, it does not allow for the effective evaluation of identification accuracies as no ground truth data is incorporated into the dataset. Low frequency measurements sampled at 1Hz from 16 different

locations within each household provide an additional layer of information to assist in the disaggregation process. However, appliances are combined into vague groupings across these 16 measurement points (e.g: 'lighting', 'miscellaneous' and 'kitchen outlets'), making it difficult to use them to approximate ground truth data. If the low frequency measurements were recorded at every individual appliance within the household, as per direct per-appliance monitoring, the dataset would be more useful for total power disaggregation technique evaluations.

AMPDS: The Almanac of Minutely Power Dataset (AMPDS) consists of energy consumption measurements taken over the period of a year in a single house, producing data that includes electricity, gas and water usage [30]. Whilst this is of some potential interest due to the combination of electrical power and other forms of energy that it contains, the low frequency of sampling (measurements taken once per minute) makes it inappropriate for use in the majority of NILM research. Certainly, the dataset is not applicable to the techniques investigated in this dissertation, especially as it lacks any form of accompanying ground truth data.

The use of a dataset compiled under real-world household conditions could be advantageous, whether used in combination with the laboratory experiment conducted in this research, or as the only source of measurements. Unfortunately, none of the datasets currently available meet the data requirements of this research, and are not appropriate for evaluating the accuracy of disaggregation techniques as they do not include accurate ground truth data. This highlights the need for high quality total power data measured from within a real-world household, with accompanying 100% accurate ground truth data.

This concludes the literature review and discussion of the NILM field as it pertains to the research area. The measurements process, including the laboratory experiment and the individual and combined appliance measurements process, are presented in Chapter 3.

Chapter 3

LABORATORY MEASUREMENTS

An overview of the measurement process is conducted, including presentation of the laboratory experiment. The device used to make the measurements is discussed, along with the electrical characteristics recorded, the choice of sampling and calculation rates and the normalisation of the resulting output. The indivual and combined appliance measurements are presented and discussed.

3.1 Measurements Overview

A series of measurements is necessary to generate total power measurements, and corresponding ground truth observations, for the implementation and comparative evaluation of TLM, CPM and EEC. Individual appliance measurements must be taken, from which libraries of exemplars can be compiled

for manual training. Combined measurements must also be taken, where appliances are operated in various combinations and sequences, creating total power measurements for the implementation of the disaggregation techniques.

To perform these measurements, eight common household appliances were assembled in the laboratory, namely; a toaster, microwave oven, sandwich maker, kettle, refrigerator, lamp, heater and fan. These appliances were connected to a multi-plug fed directly from the laboratory mains at the national standard of 230V and 50Hz. A single power measurement device was installed between the multi-plug and the wall socket, such that total power could be measured in accordance with standard NILM practice. A visual overview of the experimental setup is presented in Figure 3.1.

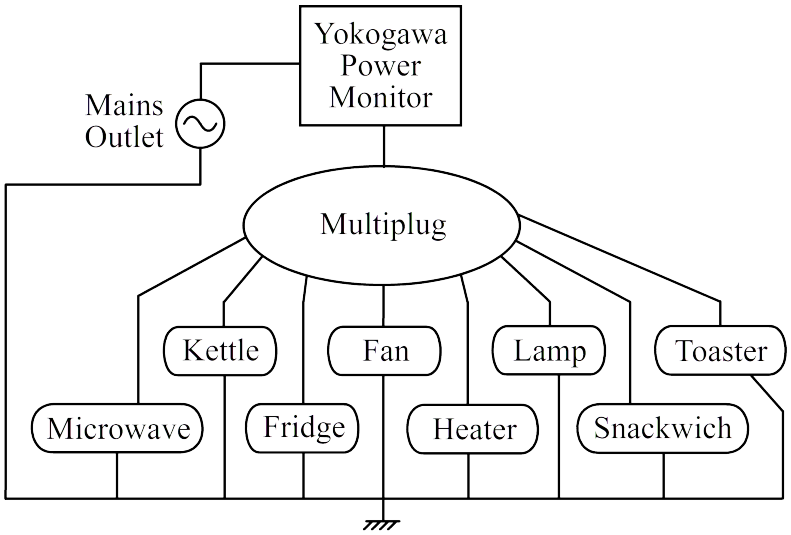

Figure 3.1: Overview of laboratory experiment.

This setup allows for measurements to be made in accordance with 100% accurate ground truth data, such that the accuracy of each technique can be properly assessed. The electrical circuit to which the laboratory experiment multi-plug was connected did not offer any isolation from the effects of other equipment operating within the building, or from variations of the

national grid. Whilst the power supplied by the national utility does fluctuate slightly in terms of voltage and frequency, initial measurements captured featured a minimal degree of noise and thus the electrical supply was deemed appropriate for the purposes of the experiment.

3.1.1 Measurements Terminology

The following terminology is used to describe the measurements throughout the remainder of the dissertation,as listed below and illustrated in Figure 3.2.

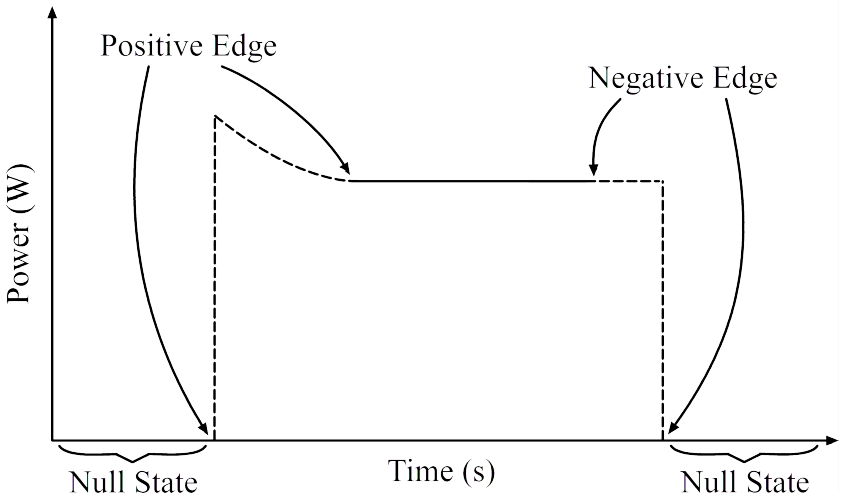

Figure 3.2: Measurements terminology presented in the context of an appliance waveform.

Positive Edge: The leading edge of the waveform, accompanied by a rapid positive change in measured power levels, where an event is detected in the total power measurement.

Negative Edge: The trailing edge of the waveform, accompanied by a rapid negative change in measured power levels, where an event is detected in the total power measurement.

Null State: A steady-state real power consumption level of approximately 0W, or 0VAR for reactive power, where no appliances are active.

The lengths of the positive and negative waveform edges, or the number of data points for which they extend away from the event edges in the power measurements, are determined by the parameters utilised for sample extraction, as discussed in Section 4.5.1.

3.1.2 Measurement Device and Sample Rate

The measurements were performed using a *Yokogawa CW240* power meter, capable of recording RMS values at up to per-cycle speeds for a 50Hz electrical system. This device was installed between the laboratory wall socket and multi-plug, as shown in Figure 3.1, such that the total power consumed by the individual and combined operation of all the appliances could be measured. The power meter measured the current, voltage and power factor, in addition to the real, reactive and apparent power at the measurement point. Of these values, the real and reactive power are the main measurements of interest for TLM, CPM and EEC.

Many NILM applications feature lower sampling rates than used in this research, with RMS values commonly reported at 1Hz (or per-second). This requires a higher sampling rate (above 1Hz) for the capture of the instantaneous voltage and current measurements, from which the RMS values are calculated. This use of a high RMS reporting rate, and hence an even higher sampling rate, allows for transients in the power waveform to be captured in greater detail than possible at lower rates. For EEC, transient features accompanying state transitions need to be captured in as much detail as possible. High sample rates are not problematic for TLM and CPM, provided that steady-states in the measurement may still be identified accurately. Thus the power meter was set to report RMS values at per-cycle intervals, as illustrated in Figure 3.3.

The *Yokogawa CW240* samples the instantaneous voltage and current signals at 128 samples per AC cycle, or 6.4kHz, using an ADC with a 16 bit resolution, as illustrated in Figure 3.3 [31]. The per-cycle RMS voltage and current

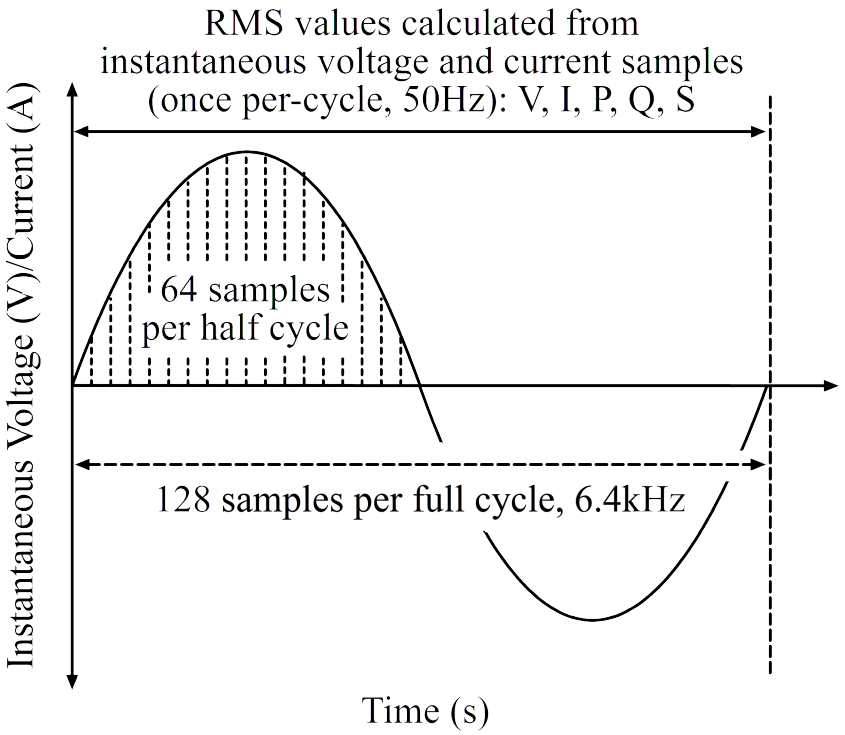

Figure 3.3: Relationship between the sampling rate and the calculated RMS measurements.

values reported by the measurement device are calculated from these samples, as shown in Equations 3.1 and 3.2, performing the required integrations at the end of each cycle. The length of a cycle is not held to a fixed value, but rather constantly adjusted to match the true frequency by monitoring zero crossings of the voltage waveform. This is preferable to using a fixed frequency value, which can lead to aliasing errors affecting the accuracy of the RMS values reported.

$$V_{rms} = \sqrt{\frac{1}{T}\int_0^T v(t)^2 dt} = \sqrt{\frac{1}{T}\sum_{t=0}^{T} v(t)^2} \tag{3.1}$$

$$I_{rms} = \sqrt{\frac{1}{T}\int_0^T i(t)^2 dt} = \sqrt{\frac{1}{T}\sum_{t=0}^{T} i(t)^2} \tag{3.2}$$

Where:

V_{rms} and I_{rms} are RMS voltage and current respectively.
$v(t)$ and $i(t)$ are the instantaneous voltage and current measurements.
T is the total number of instantaneous measurements that make up a single cycle of the AC waveform.

A combination of RMS and instantaneous values are utilised to calculate the real and apparent power values, as shown in Equations 3.3 and 3.4. Reactive power is calculated using real and apparent power, according to the power triangle, as shown in Equation 3.5. This is a commonly employed approach for accurately measuring the absolute value of reactive power, which requires that polarity be determined separately using Equation 3.6.

$$P = \frac{1}{T}\int_0^T \{v(t) \times i(t)\} dt = \frac{1}{T}\sum_{t=0}^{T} \{v(t) \times i(t)\} \tag{3.3}$$

$$S = V_{rms} \times I_{rms} \tag{3.4}$$

$$Q = \sqrt{S^2 - P^2} \tag{3.5}$$

$$Q = \frac{1}{T}\int_0^T \{v(t) \times i(t + \frac{T}{4})\} dt = \frac{1}{T}\sum_{t=0}^{T} \{v(t) \times i(t + \frac{T}{4})\} \tag{3.6}$$

Where:

P, Q and S are real, reactive and apparent RMS power respectively.
V_{rms} and I_{rms} are RMS voltage and current respectively.
$v(t)$ and $i(t)$ are the instantaneous voltage and current measurements.
T is the total number of instantaneous measurements that make up a single cycle of the AC waveform.

The alternative reactive power calculation shown in Equation 3.6 is less accurate than Equation 3.5, but does return the polarity with sufficient accuracy for it to be assigned to the reactive power value obtained from the power triangle approach.

3.1.3 Normalisation of Measurements

The voltages provided by power utilities are not necessarily constant in value, as is confirmed by the voltages measured in the laboratory experiment. In South Africa, the voltage is allowed to fluctuate within a range 10% to either side of the declared 230V level [32], [33]. These variations are aggravated by the operation of loads connected at other points within the electrical network, which can cause drops in voltage when switching events occur [34]. As the instantaneous voltage measurements affect the calculated RMS values, normalisation was universally applied to the RMS current and power measurements using Equation 3.7.

$$Y_n = \left(\frac{230}{V_m}\right)^2 (Y_m) \tag{3.7}$$

Where:

Y_n is the normalised quantity.
Y_m is the measured quantity.
V_m is the measured voltage at the time of interest.

3.2 Individual Measurements

The laboratory experiment contained eight appliances that are commonly found in households. These appliances are listed below, along with a brief description of their features and the manner in which they were operated during the experiment:

1. **Two-Slice Toaster**, rated 800W, hereafter referred to as 'toaster'.
2. **'Whistling' Kettle**, rated 2000W, hereafter referred to as 'kettle'.
3. **Desk Lamp**, rated 60W, hereafter referred to as 'lamp'.
4. **Oil-Filled Radiator Heater**, rated 1500W, hereafter referred to as 'heater'.
5. **Sandwich Maker**, rated 700W, hereafter referred to as 'snackwich'.
6. **30cm Desk Fan**, rated 35W, hereafter referred to as 'fan'.
7. **Microwave Oven**, rated 1200W, hereafter referred to as 'microwave'.
8. **Refrigerator**, rated 150W, hereafter referred to as 'fridge'.

The majority of these appliances fall into the resistive appliance category, as discussed in Section 2.6, namely; the toaster, kettle, lamp, heater and snackwich. The fan falls into the motor-driven category, the microwave-oven into the electronic power control category and the fridge into the pump category.

Each appliance was operated on its own with the power meter activated, such that individual measurements could be produced for manual training. The measurements obtained from this process, and the appliance characteristics that they reveal, are presented in Sections 3.2.1 to 3.2.8.

TLM and CPM utilise steady-state power levels in order to disaggregate the total power measurements, as discussed in Chapter 4. The real and reactive power steady-state levels extracted from the individual appliance measurements, using the technique presented in Section 4.2.2, are shown in Table 3.1.

Table 3.1: Real and reactive power steady-state levels for each appliance.

Appliance	**Real Power (W)**	**Reactive Power (VAR)**
Toaster	731.31	–
Kettle	2002.39	–
Lamp	58.05	–
Heater-Low	522.00	–
Heater-Medium	772.68	–
Heater-High	1291.06	–
Snackwich	689.39	–
Fan-Low	25.81	–
Fan-Medium	29.65	–
Fan-High	37.56	–
Microwave-Front	205.81	654.34
Microwave-Back	1063.30	384.41
Microwave-Inactive	38.73	31.39
Fridge	119.07	155.69

Table 3.1 shows that many potential appliance steady-state combinations will have real power levels that are similar in value. For example, the toaster and heater-medium states. Or the combination of the snackwich and fan-high states and the toaster on its own. Such similarities between steady-states means that slight variations in measured power levels will cause appliances, or combinations thereof, to resemble one another. This could prove problematic when attempting to disaggregate the measurements using steady-state techniques such as TLM and CPM. Due to this consideration, the choice of appliances included in total power disaggregation experiments can have a marked effect on the accuracies exhibited by steady-state disaggregation techniques.

This factor can also affect the accuracy of direct transient feature comparison NILM techniques, such as EEC, where two edges may appear more alike in shape if they possess similar levels. The degree to which this affects the disaggregation process is dependent upon the position relative to the edge events that each sample begins, as discussed in Section 4.5. In contrast to the real power steady-states, the non-zero reactive power steady-state

levels exhibit less potential for similarity and thus should be less subject to error when CPM is applied to total power measurements that contain the microwave or fridge.

3.2.1 Two Slice Toaster ('Toaster')

The toaster features a 'level' knob, which is used to control the degree to which the bread is toasted. This is a continuous control that would be expected to have no effect on the power consumption levels, affecting only the duration of the toasting process. To verify this, initial measurements were performed at two toasting levels, 50% and 100%. These were found to be identical in all respects other than the duration of their activity. Thus the cooking level selected during the performance of the experiments is of no consequence. The toaster was manually activated and deactivated via the use of a switch located on its side. The shape of the real power waveform is shown in Figure 3.4.

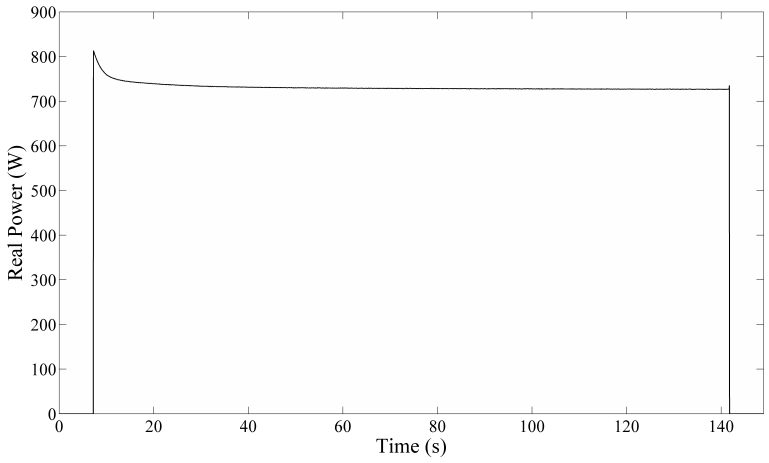

Figure 3.4: Real power waveform for toaster.

Figure 3.4 shows that maximum power is consumed upon activation, followed by a 5s curved ramp that leads down to a steady-state. The negative edge of the waveform is approximately square, with no significant features. The

reactive power plot, contained in Appendix A.1, shows that the toaster is purely resistive, the only reactive power features being the inductive transients accompanying the appliance state switching events.

3.2.2 Whistling Kettle ('Kettle')

This appliance is a very simple stainless steel kettle of 1.5 litre capacity. There is no switch built into the appliance, instead the kettle is activated or deactivated by plugging in and out its power cord. This plug insertion action can produce a transient peak in power consumption when the connection is made, due to inrush current. When the water within the kettle approaches boiling point, the device begins to emit a whistling sound from its spout. At this point the user must physically unplug it from mains power in order to stop the boiling process, as the kettle features no thermostat. The shape of the real power waveform is shown in Figure 3.5.

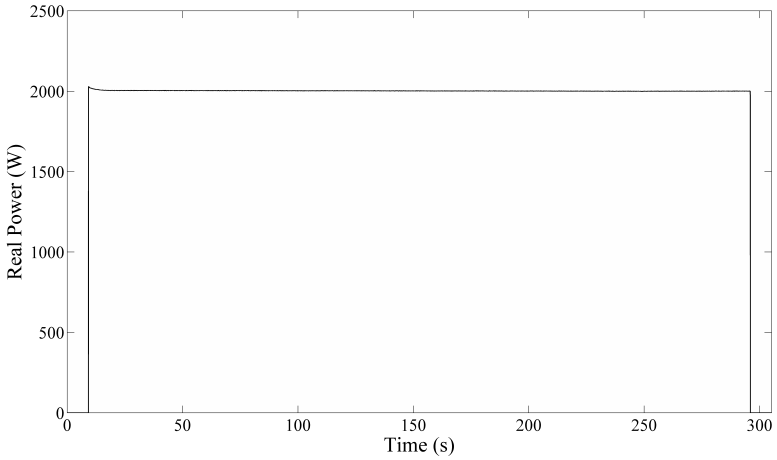

Figure 3.5: Real power waveform for kettle.

Figure 3.5 shows that maximum power consumption occurs when the kettle is initially plugged in, followed by a 10s ramp descending to a steady-state that is sustained until the user unplugs the kettle. No significant features accompany the negative edge of the waveform. The reactive power plot,

contained in Appendix A.2, shows that the kettle is purely resistive with inductive transients accompanying the plugging in and out of the cord.

3.2.3 Desk Lamp ('Lamp')

This appliance is a simple desktop lamp containing a 60W bulb. A standard contact switch is installed along the power cord, providing the means by which the lamp is activated and deactivated. The bulb is a bayonet fixture of the incandescent tungsten type. The shape of the real power waveform is shown in Figure 3.6.

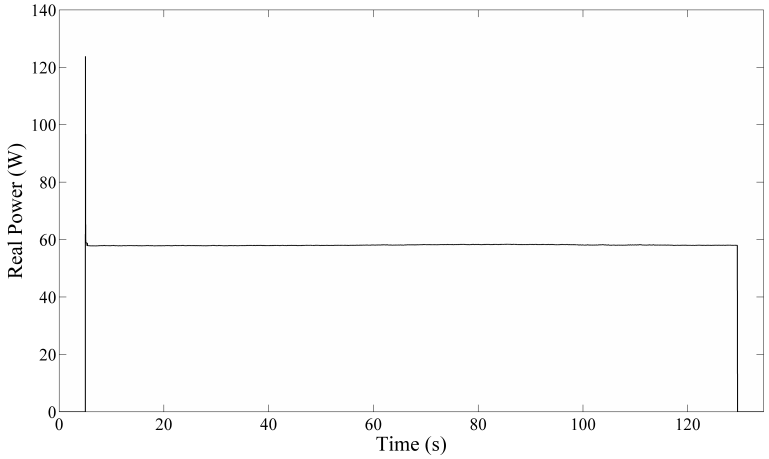

Figure 3.6: Real power waveform for lamp.

A transient power consumption peak reaching 200% of the steady-state power level may be seen when the lamp is switched on, reflecting the current flow into the filament to heat it up. The subsequent steady-state is reached within 0.5s of the appliance being activated. This remains constant until the negative waveform edge, which exhibits no significant transient features. The reactive power plot, contained in Appendix A.3, shows that the lamp is purely resistive with capacitive transients accompanying state switching events, most notably the positive edge.

3.2.4 Oil-Filled Radiator Heater ('Heater')

The heater features three heat settings, low, medium and high, each of which is selected using dedicated switches. Each setting draws significantly different levels of power, and thus the settings are treated as separate appliance states. In addition to these settings, a rotational dial is used to make finer temperature adjustments. This dial was set to 50% for all of the measurements, so that it would be consistently placed for each of the heat settings, low to high. A thermostat built into the heater controls the element by activating and deactivating it in order to maintain the selected temperature. The shape of the real power waveform is shown in figures 3.7 to 3.9 for all three heat settings, which will hereafter be referred to as 'heater-low', 'heater-medium' and 'heater-high'.

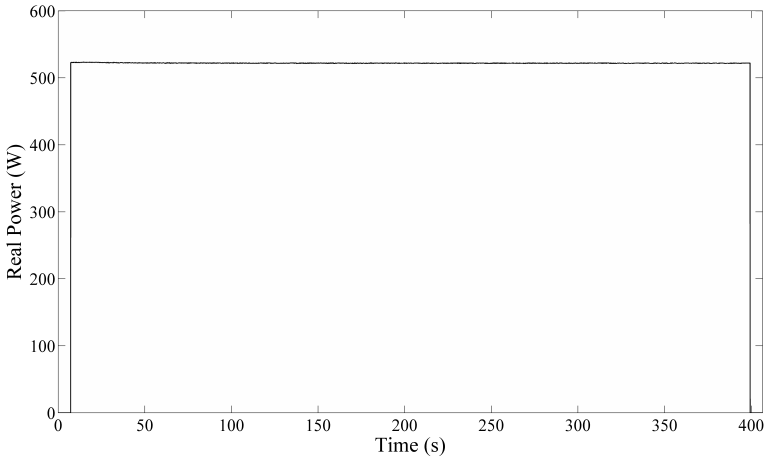

Figure 3.7: Real power waveform for heater-low.

The real power waveform shows approximately square positive and negative edges for all three heat settings, with no significant transients. Some noise may be seen in the steady-state consumption, but this is not of a level that it might be described as a specific feature such as an oscillation or ripple. The reactive power plots, contained in Appendices A.4 through A.6, show that the appliance is purely resistive for all three of the heat settings, with inductive and capacitive transients accompanying state switching events.

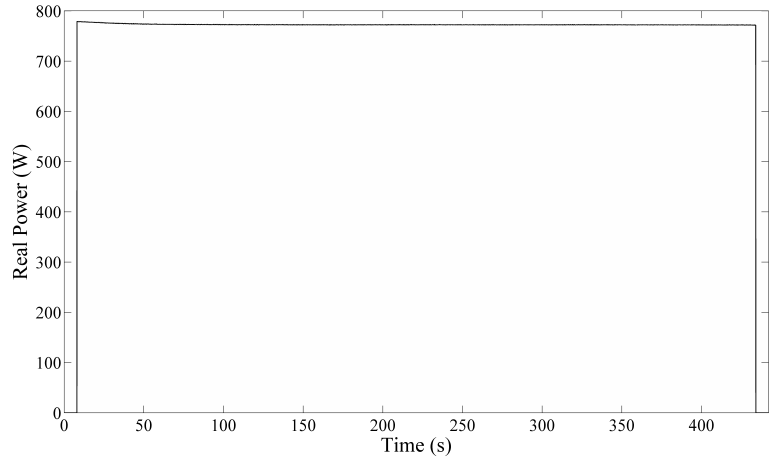

Figure 3.8: Real power waveform for heater-medium.

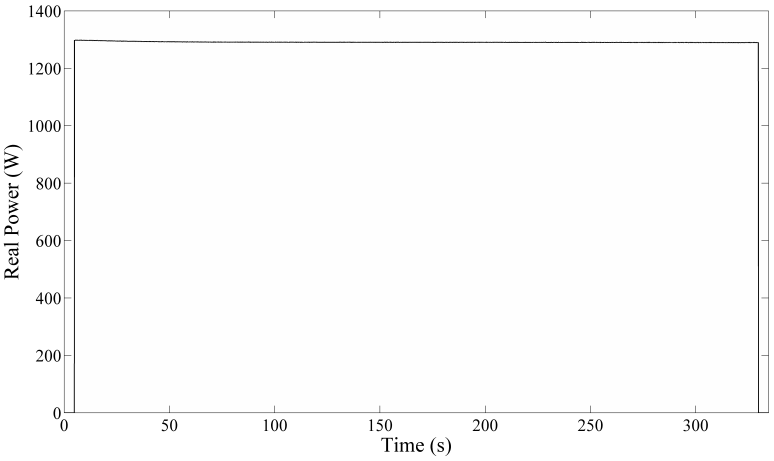

Figure 3.9: Real power waveform for heater-high.

3.2.5 Sandwich Maker ('Snackwich')

The snackwich is an enclosed toasting appliance with a casing that is heated during operation, sealing the bread as it cooks. It contains a thermostat that informs the user when the appliance has reached the optimal toasting temperature, which is signalled using green and red LEDs. Once this level is reached, current flow to the element is suspended until the temperature of the casing has dropped to a specified minimum level and the element may

be engaged again. The shape of the real power waveform is shown in Figure 3.10.

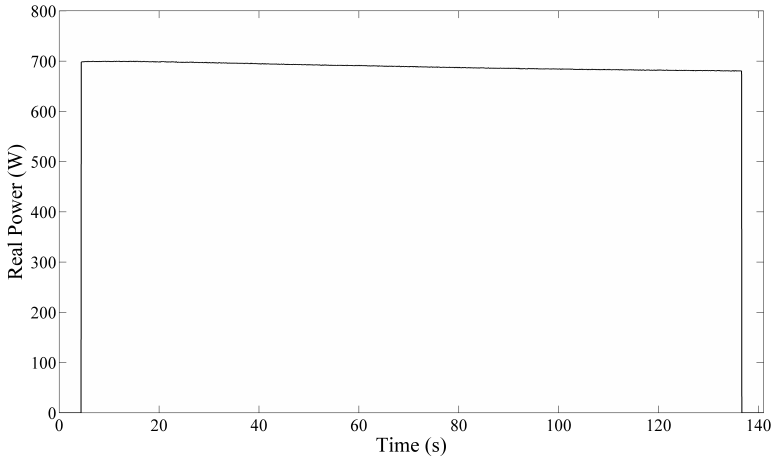

Figure 3.10: Real power waveform for snackwich.

The snackwich never reaches a steady-state condition, a constant descending ramp being found from the point of maximum power consumption at the positive edge until either the user or thermostat disengages the element. This means that the appliance must be assigned the mean of power values measured between each edge in order to approximate a steady-state, which is required for TLM and CPM to be applied to measurements in which it is incorporated. The reactive power plot, contained in Appendix A.7, shows that the appliance is purely resistive with inductive transients accompanying state switching events.

3.2.6 30cm Desk Fan ('Fan')

This appliance is a standard portable fan intended for personal cooling. It contains two DC motors, one used to spin the blades and another to rotate the head through a 60° arc. However, the head was secured in a single position during all measurements for the sake of consistency. The fan features three speed settings, low, medium and high, each of which draws a different power

level. Thus each speed settings is treated as a separate state. The shape of the real power waveforms are shown in figures 3.11 to 3.13 for all three speed settings, which will hereafter be referred to as 'fan-low', 'fan-medium' and 'fan-high'.

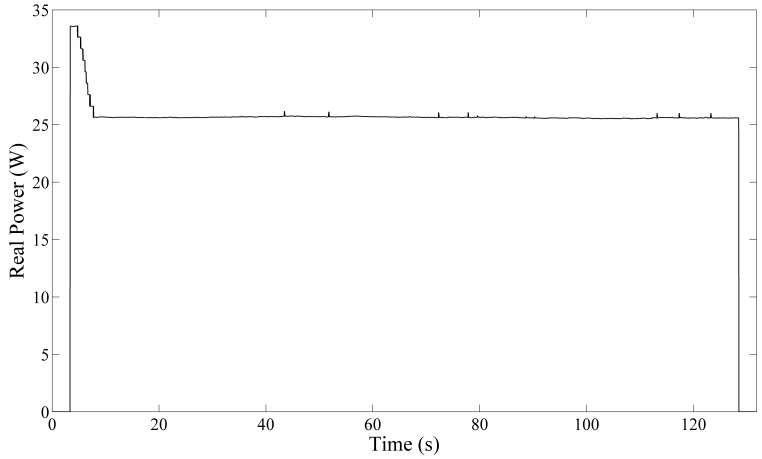

Figure 3.11: Real power waveform for fan-low.

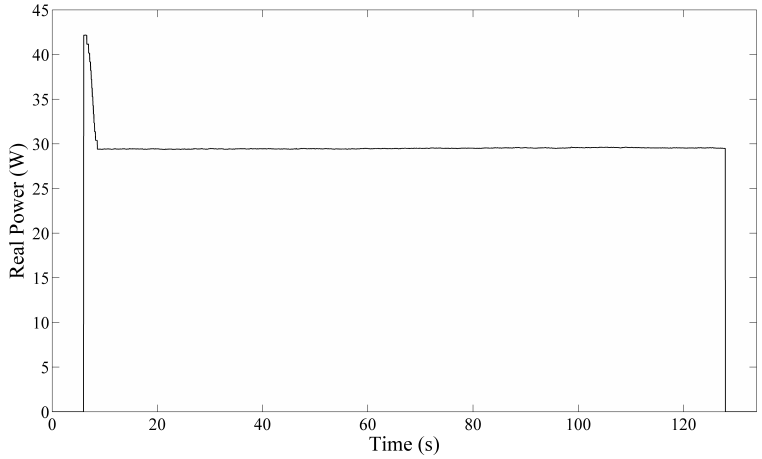

Figure 3.12: Real power waveform for fan-medium.

The real power waveforms for each of the speed settings feature similar shapes, all of which begin with a pronounced transient peak that descends to a steady-state within 4s, 2.5s and 1.5s for fan-low, fan-medium and fan-high respectively. The negative edges of the waveforms exhibit no significant

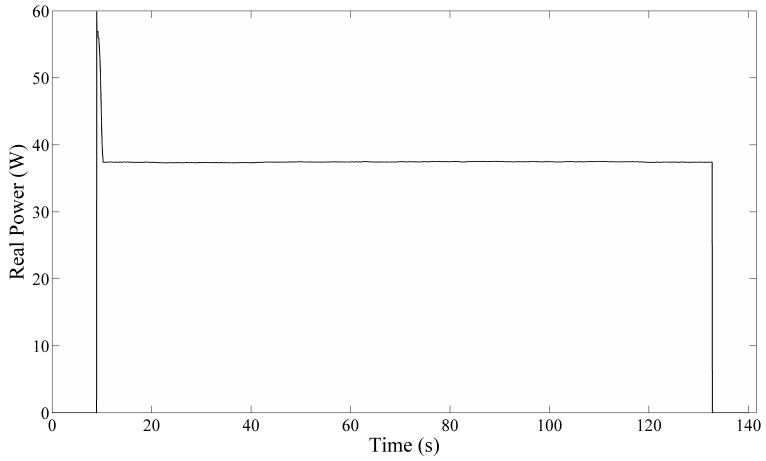

Figure 3.13: Real power waveform for fan-high.

transient features. As might be expected, the higher the fan speed setting, the greater the power consumed. The reactive power plots, contained in Appendices A.8 through A.10, show that the appliance is purely resistive during operation with capacitive and inductive transients accompanying state switching events. A series of inductive transients with 4VAR peaks may be seen during operation of the highest fan setting, particularly during the first 70s. This transient activity is reactive power noise; being too transient, and of too low a magnitude, to constitute a reactive power component for this appliance state.

3.2.7 Microwave Oven ('Microwave')

The microwave used in this experiment is a simple appliance, featuring no additional functionality beyond basic cooking. A rotational timer dial is used to set the cooking time and to begin microwave operations. A glass plate contained within the appliance cavity rotates steadily during cooking to aid in consistent heating of the meal. Five cooking settings are available, ranging from low to high. However, in reality the appliance only has a single operational state, during which the magnatron is active and a set

level of power consumed. The highest cooking level activates the magnatron throughout the cooking period, whilst each lower setting alters the duty cycle of its activation accordingly. This means that microwave operation may be characterised by a single steady-state level and set of edges, all of which represent activation of the magnatron, regardless of the chosen duty-cycle. The real power waveform of the microwave is shown in Figure 3.14 for a series of magnatron activations, as expected to be found during any cooking level below the highest.

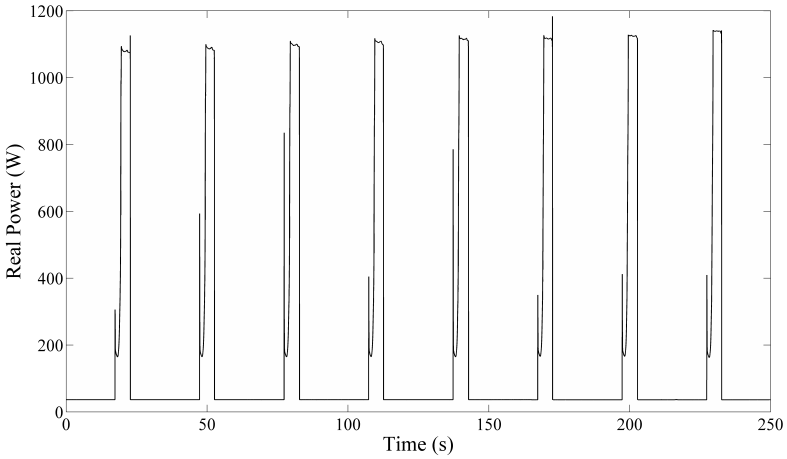

Figure 3.14: Real power waveform for microwave low power setting.

The real power waveform shows a two stage positive edge. A transient peak of inconsistent magnitude, ranging between 300W and 850W, is followed by a 1s ramp that descends from the transient peak to the 165W level. From this point, a 1s ascending ramp rises 900W to reach a noisy steady-state, 3s in length, which contains a series of ripples throughout. The negative waveform edges are approximately square, not being characterised by any significant transient features. The reactive power plot, contained in Appendix A.11, shows that the magnatron operations are not purely resistive, but have a strong reactive power component. Each magnatron activation event draws upon both real and reactive power across the same period, as shown in Figure 3.15.

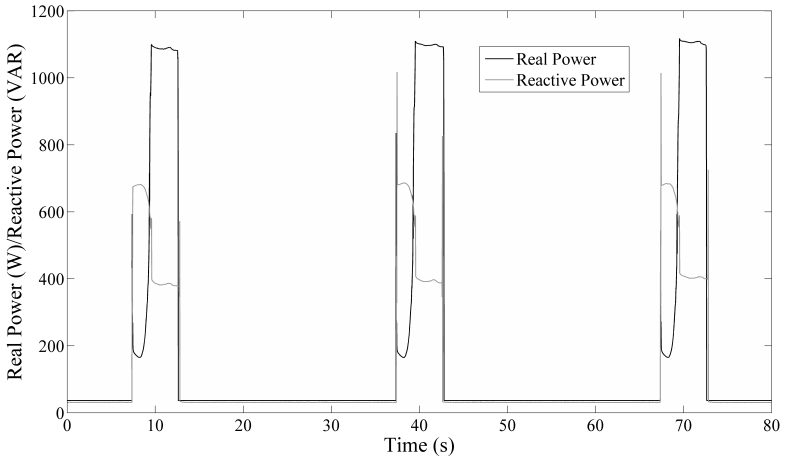

Figure 3.15: Real and reactive power waveforms for microwave low power cooking setting.

Figure 3.15 shows variation between the waveforms produced by each activation of the magnatron. Thus an average waveform must be generated that best represents this event, from which exemplar steady-state power levels and edges may be extracted. To achieve this, individual measurements of the microwave were conducted at the low, medium and high cooking levels, the results of which were averaged into a single waveform, as discussed in Section 4.5.2. This process was only followed for the real power component, as reactive power transients are not processed by TLM, CPM or EEC. However, an average of the magnatron reactive power steady-state levels was formulated for use with CPM, which requires this data.

The positive edges of the magnatron activation events consist of two distinct sections that may be considered as separate states. Splitting the positive edges into two sections assists the edge detection algorithm in accurately detecting the magnatron activation events, as discussed in Chapter 4. Both front sections of an averaged microwave exemplar waveform are shown on the same plot in Figure 3.16, where the boundary between the two is given by the vertical line. Figure 3.17 shows the entire magnatron waveform and the steady-states that correspond to each section.

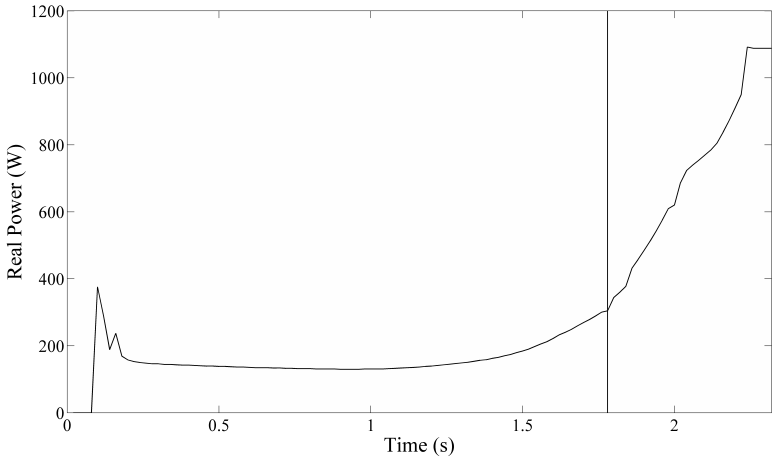

Figure 3.16: Microwave waveform front sections.

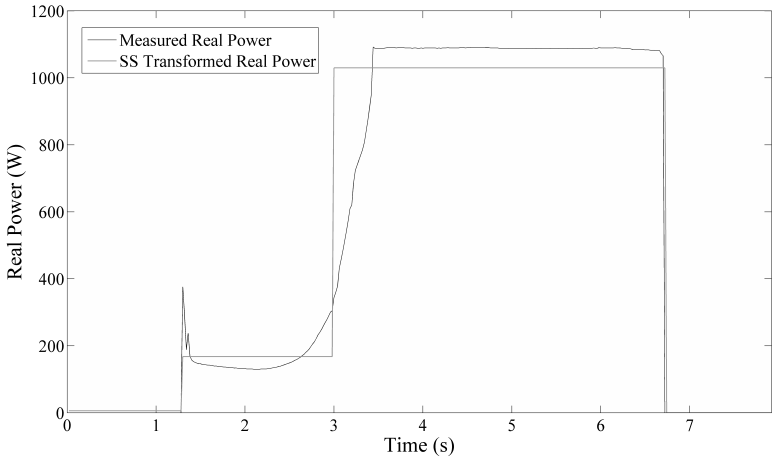

Figure 3.17: Microwave magnatron event exemplar showing the levels produced by the steady-state transformation process, namely; 'Front', 'Back' and 'Inactive'.

The microwave features an additional non-zero power consumption state that may be found where the appliance is operating, but the magnatron is not activated. This is due to a combination of electronic activity controlling the magnatron and the rotation of both the timer dial and rotating plate. It would be expected that the weight of the meal being cooked might affect the power consumed during this state, due to the size of the load on the

motor driving the rotation of the plate. For the sake of consistency, the microwave was run without a cooking load for all measurements. The real power waveform exemplar for this additional state, which was generated using the same approach used to generate the magnatron exemplar, is shown in Figure 3.18.

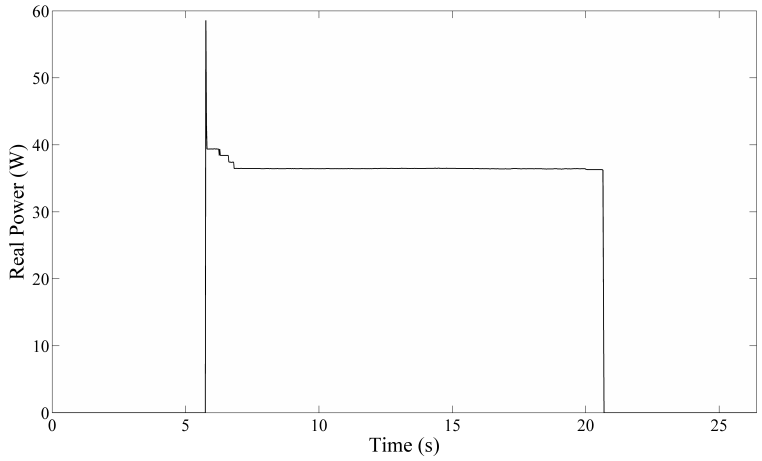

Figure 3.18: Real power waveform exemplar for microwave-inactive.

The real power waveform shows no notable transient features on either the positive or negative edge, with a consistent steady-state found between the two. The reactive power plot contained in Appendix A.12 shows that this state has a significant reactive component of approximately 31VAR.

Due to the nature of the magnatron waveform positive edges and the existence of an intermediate operational state, three appliance states which must be considered for the microwave (see Figure 3.17), namely:

Front: This refers to the section of waveform that lies between the first and second magnatron positive edges.

Back: This refers to the section of waveform that lies between the second magnatron positive edge and the magnatron negative edge.

Inactive: This refers to the section of waveform found when the magnatron is not active but the microwave is still operational. This often lies between

a magnatron negative edge and the next magnatron positive edge. However, in some cases this state will start or end the measurement, as the microwave does not always energise the magnatron immediately upon activation of the appliance.

These three states will hereafter be referred to as 'microwave-front', 'microwave-back' and 'microwave-inactive'. They are utilised with TLM, CPM and EEC, and apply to both the real and reactive power waveforms that accompany microwave operations.

3.2.8 Refrigerator ('Fridge')

This appliance is a rudimentary combination freezer and refrigerator, featuring no additional functionality such as an ice-maker or internal light. It consumes a single level of power when active, not being equipped with a thermostat to manage its temperature by altering the degree of cooling applied. Initial measurements were performed over extended periods of time in order to confirm the manner in which the fridge operates, it being divergent from that of many other similar appliances. The fridge has no switches to facilitate its activation or deactivation, and hence must be controlled by the user at the wall socket level. The shape of the real power waveform is shown in Figure 3.19.

A transient peak may be seen at the positive edge of the real power waveform, arising from the inrush current that occurs when the appliance is plugged into the mains electricity. Within 0.1s, a noisy steady-state is reached and maintained until the user disconnects the fridge from the electricity supply. The magnitude of the initial transient peak varied considerably in repeated measurements, but is always present when the fridge is activated. Sampling rates may have a large effect upon the measured values of these sorts of transients, as slower rates may miss the peak values for some iterations and catch them on others. In this research, the measurement device is sampling

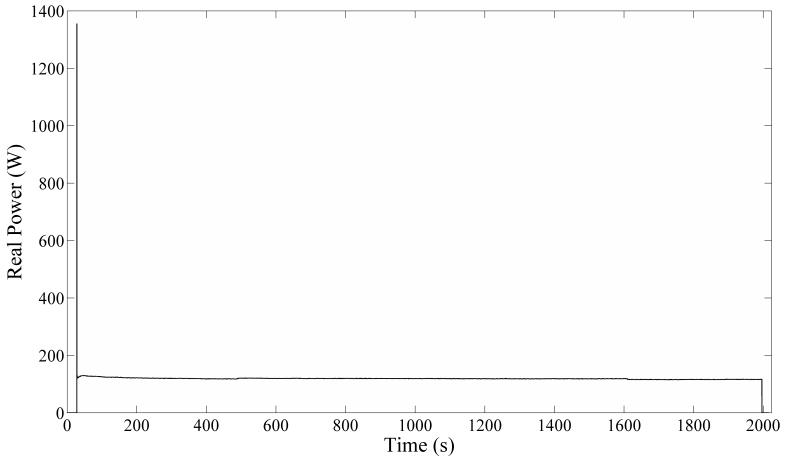

Figure 3.19: Real power waveform for fridge.

at 6.4kHz, as discussed in Section 3.1.2. This means that for each transient peak produced by the fridge, 640 samples are taken. These samples are used to produce five RMS values covering the duration of the transient, and thus the reported values may be considered accurate. The reactive power plot contained in Appendix A.13 shows that the fridge is not purely resistive, featuring an inductive steady-state of approximately 160VAR during operation. Apart from an inductive transient accompanying activation of the fridge, no significant features may be seen in the reactive power waveform.

This concludes the presentation and discussion of the individual measurements taken for each appliance. With this phase of the measurement process completed, combined total power measurements featuring combinations of appliance operations could be generated, as presented in Section 3.3 below.

3.3 Combined Measurements

In order to evaluate TLM, CPM and EEC, total power measurements featuring various combinations of the appliances presented in Section 3.2 were made. Ground truth data schedules were created for nine experiments,

featuring combinations of appliance activities that might realistically occur within a real-world household. The ground truth data was rigorously enacted using the laboratory experiment. Due to human error, slight discrepancies may be found between the timing of events in the ground truth data and those detected in the measurements. These timing discrepancies, in the order of one to two seconds, are small enough to be of no consequence when evaluating the accuracy of TLM, CPM and EEC.

Where appliances are equipped with a variable selector, the related parameter only serves to affect the time of engagement or duty-cycle. This is the case for the toaster, heater and microwave, as discussed in Section 3.2 Thus, each appliance referenced in the ground truth data could theoretically be operated with any combination of settings without compromising the applicability of the individually extracted exemplars to the total power measurements. However, the same appliance settings used to generate the exemplars were used throughout all nine of the combined measurements, in accordance with good experimental practice.

Each combined measurement contains a varied number of appliances, from a minimum of two, up to a maximum of four. However, where multi-state appliances are involved, the availability of multiple operational states for a single appliance allows for additional complexity to be introduced into the total power measurements. The degree of complexity reached with even such low numbers of appliances is sufficient to severely test the disaggregation techniques, as demonstrated by the results in Chapter 5. Thus it was deemed unnecessary to perform further combined measurements with larger numbers of appliances.

The combined measurements are briefly presented and discussed in sections 3.3.1 to 3.3.9 below. The reactive power plots for the combined measurements 1 through 9 are contained in Appendices A.14 to A.22 respectively. The tables of ground truth data corresponding to each combined measurement may be found in Appendices B.1 to B.9.

3.3.1 Combined Measurement 1: Fan and Toaster

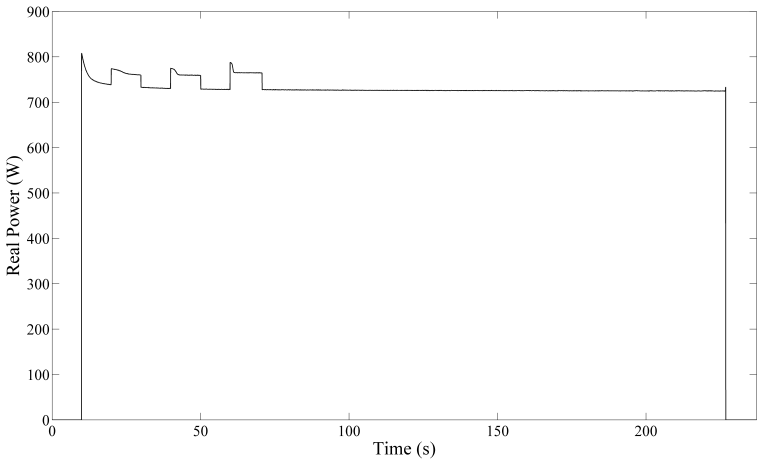

Figure 3.20: Combined Measurement 1 real power plot.

The real power waveform shown in Figure 3.20 shows the overriding power consumption of the toaster, which is 1824% greater than that of the fan-high state. Whilst each of the three fan settings feature different levels of power consumption, these differences appear minimal in the context of the toaster waveform. The fan-low, fan-medium and fan-high waveform positive edges may be found at 20s, 40s and 60s respectively. As is consistent with the individual toaster measurements discussed in Section 3.2.1, a descending ramp may be found between the toaster and fan-low waveform positive edges found at 10s and 20s respectively. The steady-state power level calculated for this section of the measurement, and for the subsequent fan-low and toaster combination, exceed the expected levels due to this ramp, introducing error into the associated TLM and CPM disaggregation accuracies.

As both appliances are purely resistive, the reactive power waveform shown in Appendix A.14 contains only inductive and capacitive transients where the toaster is switched on and off, as is consistent with the individual appliance measurements.

3.3.2 Combined Measurement 2: Toaster and Microwave

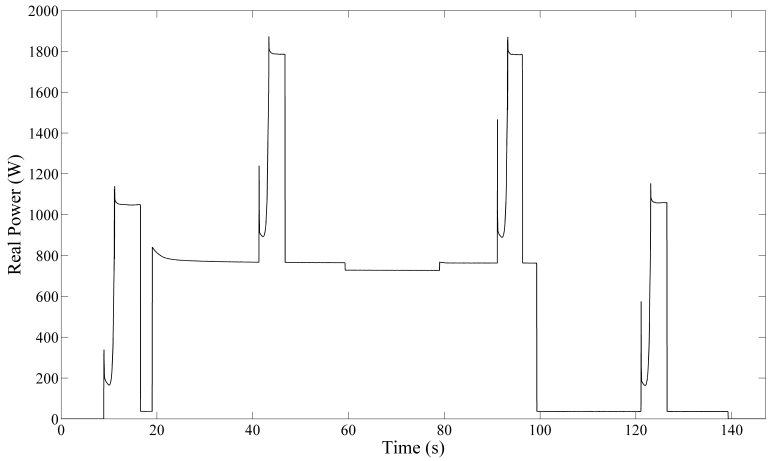

Figure 3.21: Combined Measurement 2 real power plot.

A series of microwave magnatron events may be seen throughout the measurement, shown in Figure 3.21, with waveform positive edges at 9s, 41s, 91s and 121s. The distinctive shape of these edges may be easily distinguished from the relatively featureless toaster waveform positive edge found at 19s. In this measurement the shapes of the magnatron event waveform edges are not significantly altered where they are found in combination with the toaster, due to the relatively square nature of the toaster waveform. The combination of two irregularly shaped waveforms will produce a combined waveform that is more difficult to relate back to the individual appliances that produced them than is the case in this measurement. A 40W drop in power consumption may be observed between 59s and 79s due to the complete deactivation of the microwave, which removes the microwave-inactive state from the measurement, leaving the toaster to operate alone for 20s.

As the microwave is the only appliance included in this measurement that has a reactive power component, it dominates the reactive power measurement shown in Appendix A.15. Inductive and capacitive transients may also be seen where state switching events occur for both appliances. Whilst the reactive power measurement is largely consistent with expectations aris-

ing from the individual microwave-state measurements, as discussed in Section 3.2.7, there are sections in the measurement where levels of approximately 0VAR are measured when a microwave-inactive state of approximately 31VAR should be measured. This occurrence may be observed in Appendix A.15 from; 19s to 41s, 47s to 59s, 79s to 91s, and 96 to 99s.

The toaster is the only other appliance active during these sections of the measurement, which would suggest that some combination of both appliances has lead to the introduction of a negative reactive power component (capacitive) into the measurements. It is known that the toaster is purely resistive, thus this decrease in reactive power must be due to some physical property of the laboratory experiment circuit that is external to the appliances themselves. Alternatively, it could be ascribed to a measurement error, perhaps due to the combination of methods used by the *Yokogawa CW240* to determine the reactive power magnitude and polarity, as discussed in Section 3.1.2.

It is expected that this type of inconsistency be found within total power measurements obtained from real-world NILM system implementations, especially where the measurement devices utilised offer inferior performance to the *Yokogawa CW240* and complex power networks are involved. Thus the presence of this error in the combined measurements presents an opportunity to test the performance of CPM under such conditions. Of the reactive power appliance states, the microwave-inactive state is the most likely to be adversely affected by such inconsistencies, as its low steady-state power level of 30VAR is the closest to another state in the reactive power appliance combination vector, namely the null state. Thus, due to its low reactive power consumption, any reductions in the total measured reactive power are more likely to lead to the microwave-inactive state being incorrectly identified as a null than for any other state, depending on the levels of reactive power reduction power being experienced. Two separate variants of CPM, focussed around the microwave-inactive state, were developed to determine whether low power appliance states should be included in the reactive power matching

stage of CPM under such conditions, as discussed in Sections 4.4 and 5.3.1.

3.3.3 Combined Measurement 3: Toaster and Microwave

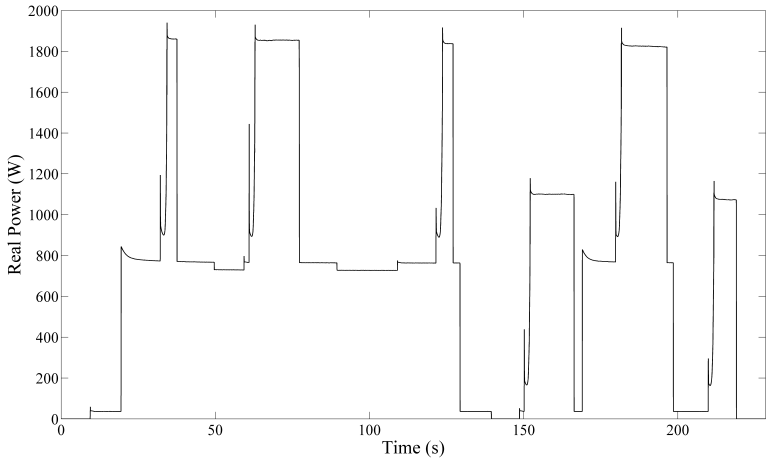

Figure 3.22: Combined Measurement 3 real power plot.

This combined measurement, shown in Figure 3.22, contains the same appliances as Combined Measurement 2. However, it features a higher degree of complexity due to the sequence of appliance operations enacted during the measurement process. Six magnatron events may be seen, with positive waveform edges at 32s, 61s, 122s, 150s, 180s and 210s. The length of the magnatron events varies noticeably in each case, due to the use of both the low and medium microwave cooking settings; where the low setting produces 6s magnatron events (e.g: from 32s to 38s) and the medium setting produces 16s magnatron events (e.g: from 61s to 77s). The last medium cooking level magnatron event, which starts at 210s, is reduced to only 9s in length due to the deactivation of the microwave at 219s. The toaster is active for two periods within the measurement, initially being operational between 19s and 129s, and then appearing again between 169s and 199s.

As with Combined Measurement 2, the reactive power components of each microwave state dominate the reactive power measurement, shown in Ap-

pendix A.16. Inductive and capacitive transients may also be seen where appliance state switching events occur. The same microwave-inactive state reactive power measurement error discussed in Section 3.3.2 may be found where the toaster and microwave-inactive state are simultaneously active. This is shown in Appendix A.16, in the sections of the reactive power measurement from; 19s to 32s, 38s to 50s, 59s to 61s, 77s to 89s, 109s to 122s, 127s to 129s, 169s to 180s and 196s to 199s.

3.3.4 Combined Measurement 4: Lamp and Snackwich

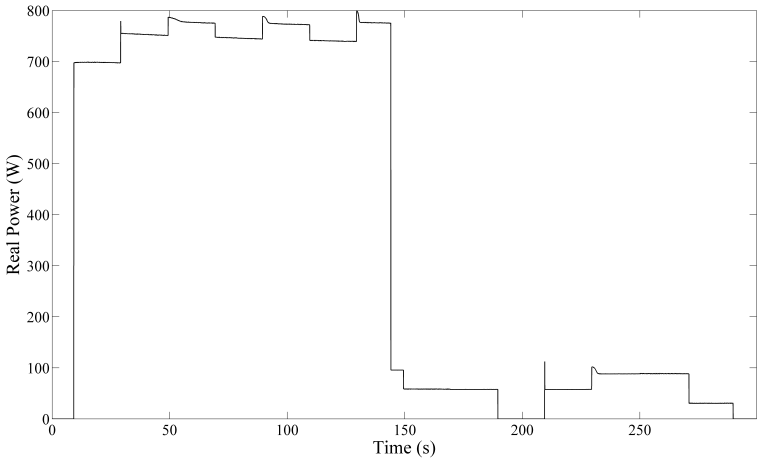

Figure 3.23: Combined Measurement 4 real power plot.

The power consumption of the snackwich is 1088% greater than that of the lamp, and 1713% greater than that of the fan-high state. Thus the operation of the snackwich, occurring between 9s and 144s, dominates the real power plot for Combined Measurement 4, as shown in Figure 3.23. The lamp is found twice within the measurement; firstly being active between 29s and 190s, and appearing again between 210s and 271s. Both of the waveform positive edges that accompany the lamp activation events feature a transient peak of 0.1s in duration, with magnitudes of 80W and 110W for the first and second events respectively.

The remainder of the waveforms visible in the measurement are due to the three fan states, with their recognisable waveform positive edges being found at 49s, 89s, 130s and 230s for the states fan-low, fan-medium, fan-high and fan-medium respectively. The negative edge of the snackwich waveform may be seen at 144s, followed by the smaller negative edges of the fan-high state waveform at 149s and the lamp waveform at 190s. The series of steps down to a 0W level of power consumption that result from this sequence of events were created by the snackwich thermostat deactivating the appliance during a period of fan-high and lamp operation. Despite being an automated state change, this event has been accurately captured in the ground truth data contained in Appendix B.

The reactive power plot for this combined measurement, shown in Appendix A.17, contains only inductive and capacitive transients, which are found where appliance state switching events occur. This is as expected, given that all three appliances included in the measurement are purely resistive.

3.3.5 Combined Measurement 5: Fan, Lamp and Heater

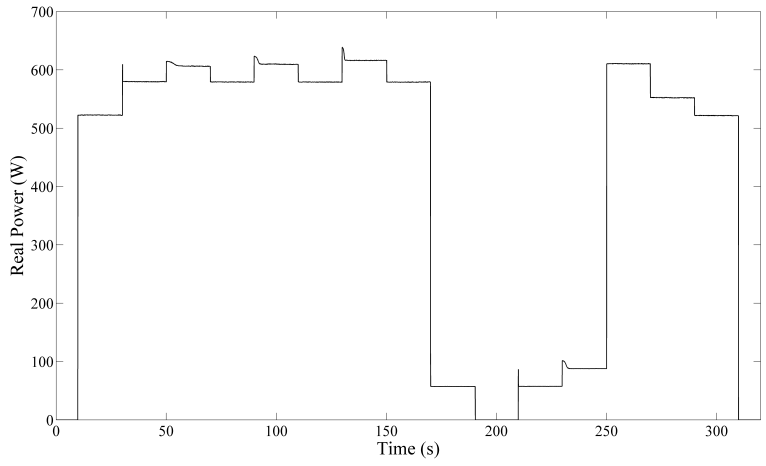

Figure 3.24: Combined Measurement 5 real power plot.

Of the three heater heat settings available, only the heater-low state was

utilised for this combined measurement. Thus only six appliance state combinations are contained within the real power plot shown in Figure 3.24, namely; heater-low, fan-low, fan-medium, fan-high, lamp and null. The expected real power consumption for the heater-low state is 1274% greater than that of the fan-high state, and 800% greater than that of the lamp. Thus the heater-low waveform is the most prevalent feature in Figure 3.24, initially being active between 10s and 170s, and appearing again between 250s and 310s. The lamp is activated twice, with positive waveform edges being found at 30s and 210s, each of which features a 90W transient of 0.1s duration. The fan is operated four times, with positive waveform edges located at 50s, 90s, 130s and 230s for the states fan-low, fan-medium, fan-high and fan-medium respectively. Where the heater is deactivated at 170s it does not return to a 0W steady-state level due to the lamp waveform, which continues until 190s.

Only transients, both inductive and capacitive, are found in the reactive power plot for Combined Measurement 5, as shown in Appendix A.18. As the heater, fan and lamp are all purely resistive, this is consistent with expectations. The transients seen in Appendix A.18 may be ascribed to appliance state switching events.

3.3.6 Combined Measurement 6: Fridge, Heater and Microwave

The first feature found within the real power plot for this combined measurement, shown in Figure 3.25, is a 1850W transient peak that accompanies the activation of the fridge at 10s. The next readily recognisable waveforms within the measurement are produced by magnatron events, with waveform positive edges found at 80s, 107s, 137s, 179s, 218s, 248s and 278s. The third appliance appearing in Figure 3.25 is the heater, which is operated on the medium and high settings. Heater-medium states may be found between 59s and 100s, and between 219s and 259s. A heater-high state may be found between 119s and 159s.

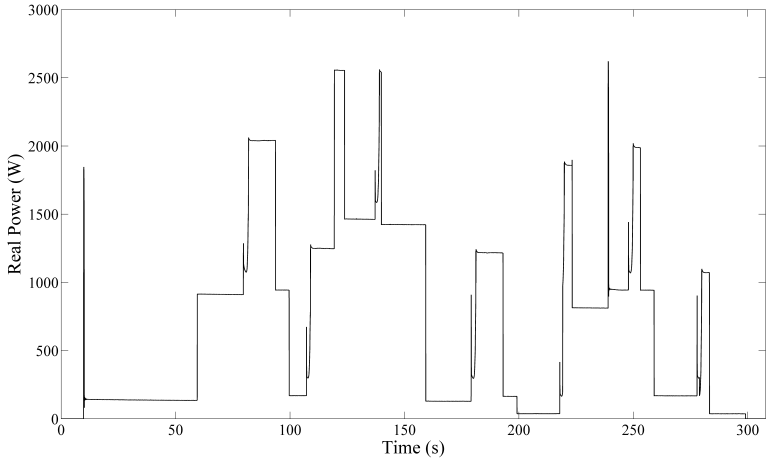

Figure 3.25: Combined Measurement 6 real power plot.

Two of the microwave magnatron events are interrupted by the operations of other appliances, splitting the waveforms into two distinct sections. The first occurs at 119s, where the heater-high state becomes active during the magnatron event found between 107s and 124s. The second occurs at 219s, where the heater-medium state interrupts the positive edge of the magnatron event waveform found between 218s and 223s. This event illustrates one of the fundamental flaws in the assumption, commonly made in NILM research, that no appliance operations occur simultaneously. Whilst it is certainly unlikely that a user activates two appliances at the same time, automatically controlled state transitions may well take place concurrently with manual ones.

Both the fridge and microwave have reactive power components which feature in the reactive power measurement, shown in Appendix A.19. In addition to the steady-state levels, inductive transients are found where appliance state switching events occur. Similarly to the reactive power measurement error discussed in Sections 3.3.2 and 3.3.3, the microwave-inactive state is affected whilst being operated concurrently with the heater. Between 124s and 137s the approximately 31VAR expected for the microwave-inactive state is not present in the measurement, where the heater-high and fridge states are also

active. Between 223s and 239s, the microwave-inactive state appears not to be present in the measurements, where the ground truth data reflects that the microwave is active concurrently with the heater-medium state. This indicates the presence of the same inconsistency observed in Combined Measurement 2, where a certain combination of appliances leads to a reduction in the measured reactive power levels. Between 140s and 159s, where the fridge is operating concurrently with the heater, a disparity of approximately 30VAR be again be seen between the measured and expected reactive power levels. This disparity disappears when the heater is deactivated at 159s. As the heater is known to be purely resistive, as presented in Section 3.2.4, this supports the discussion conducted in Section 3.3.2. Further variations on this inconsistency may be found throughout the reactive power measurement.

3.3.7 Combined Measurement 7: Lamp, Heater and Toaster

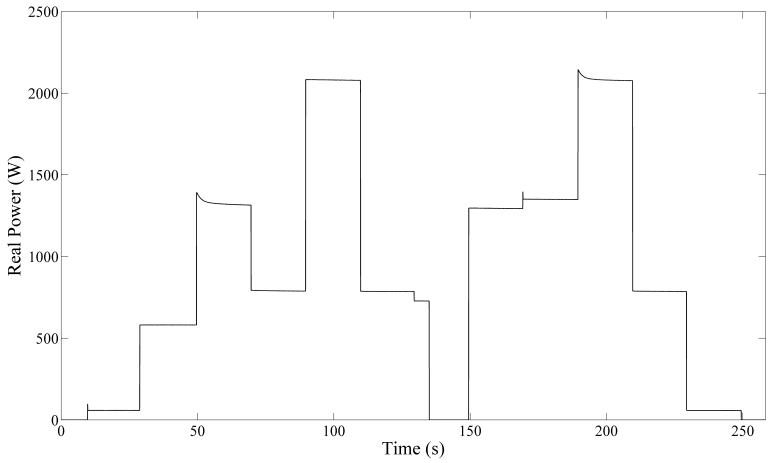

Figure 3.26: Combined Measurement 7 real power plot.

All three heat setting states for the heater are included in this combined real power measurement, shown in Figure 3.26. A heater-low state may be found between 29s and 70s, a heater-medium state between 90s and 100s and

a heater-high state between 149s and 210s. The toaster waveform positive edges, with the readily recognisable ramp from peak to steady-state discussed in Section 3.2.1, may be found at 50s and 190s. The lamp is activated twice, being operational between 10s and 129s, and between 169s and 249s. Apart from a 100W transient of 0.1s in duration that accompanies the lamp waveform positive edge at 169s, the toaster is the only appliance to exhibit a distinctive transient feature on its waveform positive edges. The rest of the appliance state transitions in Figure 3.26 are square in shape, with no distinctive transient features.

As all three of the appliances are purely resistive, the reactive power measurement contains only inductive and capacitive transients, as shown in Appendix A.20. These are found wherever appliance state transitions occur.

3.3.8 Combined Measurement 8: Fridge, Kettle and Lamp

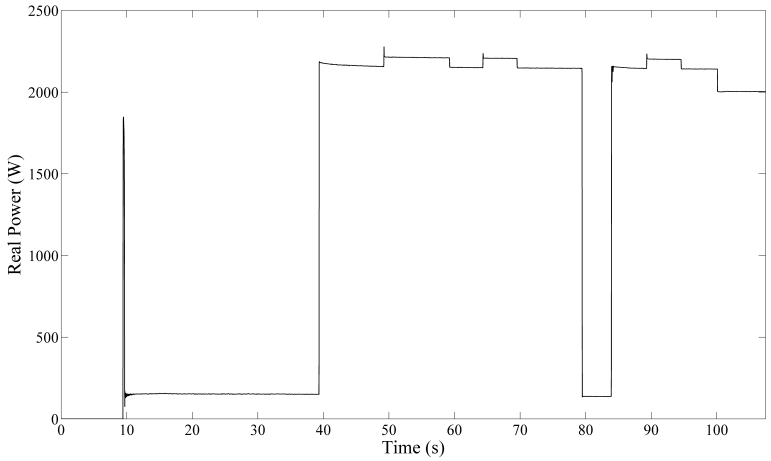

Figure 3.27: Combined Measurement 8 real power plot.

Similarly to Combined Measurement 6, the positive edge of the fridge waveform found at 9s is the first significant feature of this combined measurement, as shown in Figure 3.27. The kettle is the highest rated power appliance of

all of those included in the laboratory experiment, consuming 1582% more power than the fridge, and 3352% more than the lamp. Where it becomes active at 39s and 84s, the power level rises by 2030W. The fridge is not deactivated until 100s, the three smaller power consumption events between 49s and 95s being due to operation of the lamp. The lamp positive waveform edges, located at 49s, 64s and 89s, exhibit the same 0.1s transients discussed in Section 3.3.7, ranging between 80W and 120W in magnitude. Unlike the other combined measurements, the final steady-state shown in Figure 3.27 is not a null state, as the kettle was not deactivated before the end of the measurement.

The fridge is the only appliance included in this combined measurement that has a reactive power component, and hence it is the sole source of the reactive power steady-states shown in Appendix A.21. The expected inductive and capacitive transients may be found accompanying appliance state switching events throughout the measurement. For all of the reactive power measurements where the kettle is active, namely from 39s to 79s and from 84s to the end of the measurement, an approximately 50VAR discrepancy between the measured and expected reactive power steady-state levels may be seen. However, this discrepancy is not seen where the kettle is deactivated between 79s and 84s. The kettle is purely resistive, as presented in Section 3.2.2, and thus this occurrence may be considered to be a variation on the same reactive power measurement inconsistency discussed in Section 3.3.2.

3.3.9 Combined Measurement 9: Toaster, Microwave, Heater and Lamp

This is the sole combined measurement to include four appliances, and thus might be expected to feature the highest levels of complexity. However, there is no point in time where all four appliances are simultaneously operational. Thus the measurement contains more variety in terms of appliance states, but it does not contain a greater complexity of appliance state combinations.

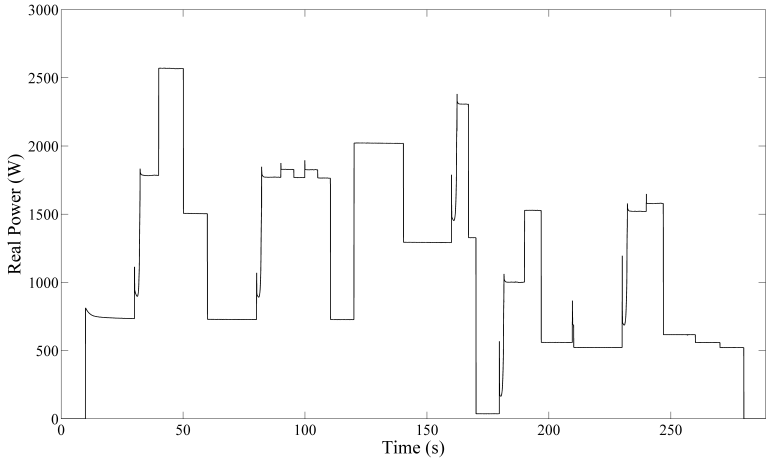

Figure 3.28: Combined Measurement 9 real power plot.

The toaster waveform positive edge, located at 10s, is the first feature to be found in the measurement, as shown in Figure 3.28. The expected ramp may be seen descending from the toaster positive edge, until it is interrupted at 30s by a microwave magnatron event. Figure 3.28 contains four further magnatron events, with waveform positive edges found at 80s, 160s, 180s and 230s.

The lamp is activated three times, being operational for two 5s periods starting at 90s and 110s, and for a 20s period starting at 240s. Each lamp waveform positive edge features a 0.1s transient, ranging between 100W and 120W in magnitude. The heater is the only appliance included in this combined measurement that does not have a significant transient accompanying its waveform positive edges. The three heater heat setting states are active during the following time periods; heater-medium between 40s and 60s, heater-high between 120s and 170s, and heater-low between 190s and 280s.

The microwave is the only appliance included in this combined measurement that has a reactive power component. Thus all of the steady-states in the reactive power measurement may be attributed to its operations, as shown in Appendix A.22. Both inductive and capacitive transients may be found

accompanying appliance state switching events. Similarly to Combined Measurement 6, the microwave-inactive state may be seen to be affected by the presence of the heater wherever the two are active at the same time, namely from; 167s to 170s, from 197s to 210s, and from 247s to 260s. For these steady-states in the measurement, the measured reactive power level is approximately 30VAR lower than expected. This discrepancy between the measured and expected reactive power levels is consistent with the observations made for the microwave-inactive and toaster appliance state combination in Sections 3.3.2 and 3.3.3. Between 210s and 230s, where the heater-low state is active alone, an unexpected 18VAR of reactive power may be found. As the heater is purely resistive, this measured reactive power level may be ascribed to noise.

This concludes the presentation of the measurements process, including discussion of both the individual and combined appliance measurements. The fundamental mechanics and software implementations of TLM, CPM and EEC are presented in Chapter 4.

Chapter 4

IMPLEMENTATION OF TECHNIQUES

An overview of the system developed for the implementation of TLM, CPM and EEC is conducted, followed by discussion of the underlying processes required in order for each of the three disagreggation techniques to be applied to the measurements. The implementations of TLM, CPM and EEC are presented, both in the form of conceptual overview and through more specific discussion of the key functions from which they are constituted. Throughout the chapter, reference is made to relevant samples of implementation code, all of which are contained in Appendix C.

4.1 System Overview

Before implementing TLM, CPM and EEC, as discussed in Sections 4.3, 4.4 and 4.5 respectively, the power measurements must be processed using the

steps presented in Section 4.2. All measurements processing and disaggregation technique implementations are performed in *MATLAB* [3].

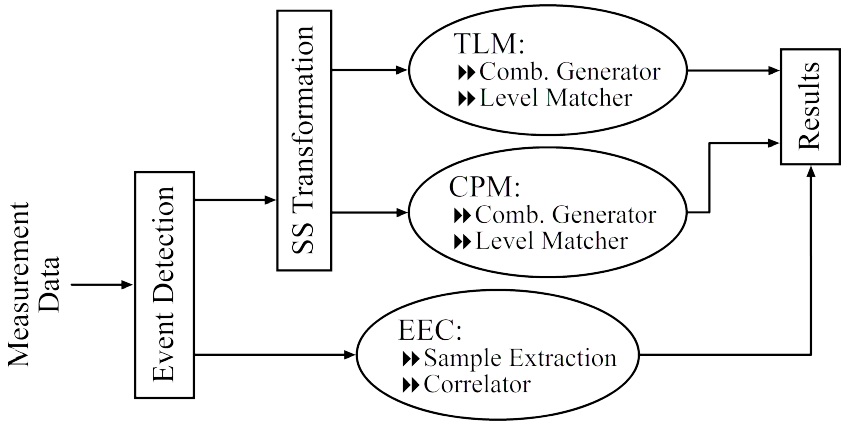

Figure 4.1: System overview showing relationship between fundamental measurements processing functions and disaggregation technique components.

Figure 4.1 provides a visual overview of the system created within *MATLAB*, showing how a few basic functions provide a platform for each of the components comprising TLM, CPM and EEC to be implemented. In the remainder of this chapter the fundamental measurements processing functions are presented, followed by the main disaggregation techniques.

4.2 Underlying Processing Approaches

The most fundamental of all the preparatory steps is the transfer of the individual and combined measurements data from the power meter into *MATLAB*, where it is captured into vectors for further use. During this process all data is normalised, using the equations presented in Section 3.1.3. All three disaggregation techniques require that every appliance state transition contained within the total power measurement be identified, although they utilise this information in different ways.

EEC uses the appliance state transition information directly, extracting samples from around edges that were detected in the measurement. However,

TLM and CPM require that the measurement be broken down further into a series of steady-states, which requires an additional measurements processing stage referred to as 'steady-state transformation' in this dissertation. The event detection and steady-state identification functions are discussed in more detail in Sections 4.2.1 and 4.2.2 respectively.

4.2.1 Event Detection

The *MATLAB* code for this function, `eventDetection`, may be seen in Appendix C.1. As illustrated in Chapter 3, every appliance state transition features a distinct edge which may be observed in both the individual and combined measurements. Whilst it may theoretically be possible for an appliance state change to be characterised by gradual ramps instead of a defined edge, this is not the case with any of the appliances included in the experiment. Even where start-up ramps are involved and such a waveform may be expected, such as for the slow acceleration from standstill of a washing machine drum shown in Figure 4.2, a marked edge still accompanies the start and end points of the state transitions.

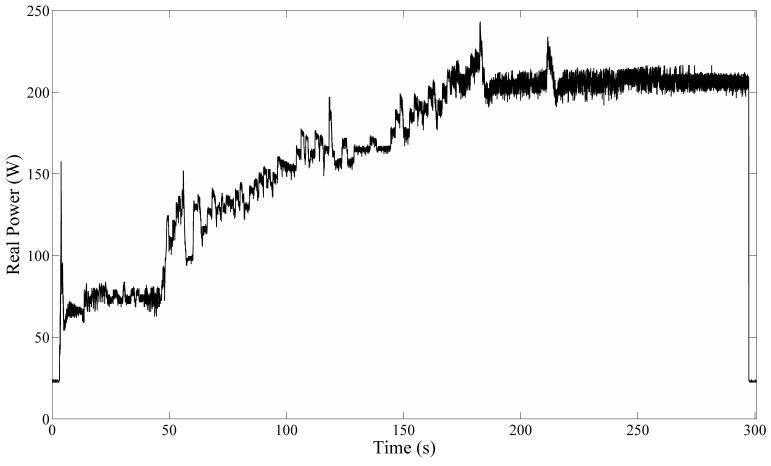

Figure 4.2: Washing machine spin cycle real power plot.

The event detection algorithm compares all adjacent data points in the mea-

surement, assigning edge status wherever a rapid change in power consumption is found that exceeds a 15W threshold. The direction of the change is also noted, so that it may be known whether the edge is constituted by a positive or negative change. This is important for EEC, where positive and negative edge samples are compared separately, as discussed in Section 4.5. To ascertain the optimum threshold level, a range of plausible options were tested until a value was found that resulted in perfect edge detection for all of the individual and combined measurements. Due to the experimental approach used to determine this threshold value, it may only be deemed to be optimal for the experiments and measurements conducted within this dissertation.

A measurement, containing several fan waveforms, which has been processed using the event detection algorithm is shown in Figure 4.3, where the direction of the triangle markers indicate the direction assigned to each detected edge.

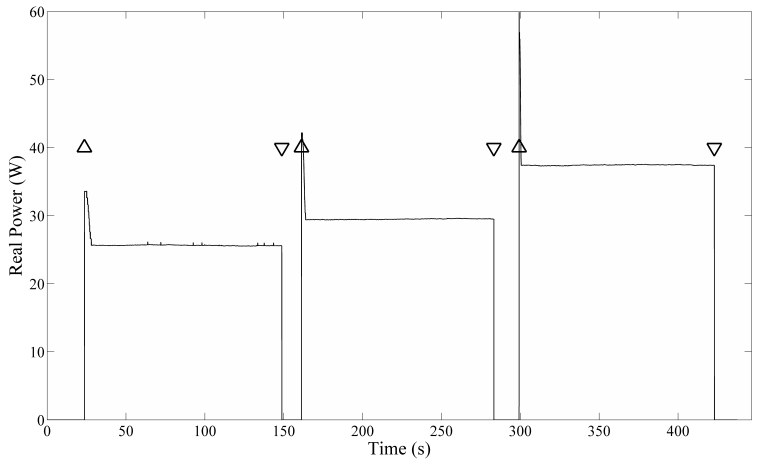

Figure 4.3: Event detection applied to multi-state fan measurement.

Many edges in the power measurements occur over a number of data points, thus the potential exists to incorrectly classify a single dramatic power change as a series of edges. To counter this problem, averaging windows and other similar smoothing techniques were initially included in the algorithm. How-

ever, these were ultimately discarded in favour of a simpler approach that yielded better results. Once an edge has been identified, the algorithm is disabled for the next 50 data points. To avoid error, this value has to be below the minimum time occurring between any two appliance state transition events across all nine combined measurements, which is 51 data points.

Event detection is only ever applied to the real power waveform. For CPM, where reactive power steady-states are required, the real power waveform is used to find the locations in the total power measurements where appliance state transitions occur. These points in time are then transposed across to the reactive power measurements, rather than performing a separate event detection pass. This is necessitated by the magnitudes of the transients that accompany appliance state switching events, as discussed in Chapter 3. These transients require that very high event detection threshold levels be set, which in turn leads to smaller appliance edges going undetected. However, the location of edges in the reactive power measurements almost exactly mirror those found in the real power measurements and thus this transposition is not problematic.

4.2.2 Steady-State Transformation

The *MATLAB* code for this function, `ssTransformation`, may be seen in Appendix C.2. The steady-state transformation process takes the edges found by the event detection function and finds the mean of the real or reactive power values lying between them, transforming the original waveform into a sequence of steps, as shown in Figure 4.4, where the grey plot shows the steady-state identification output. This transformation of the power measurements is required by both TLM and CPM, where the series of steady-states that comes to represent each of the combined measurement must be compared against possible combinations of power levels.

In some cases, the transient features within an appliance waveform lead to the calculation of mean values that do not match the expected steady-state

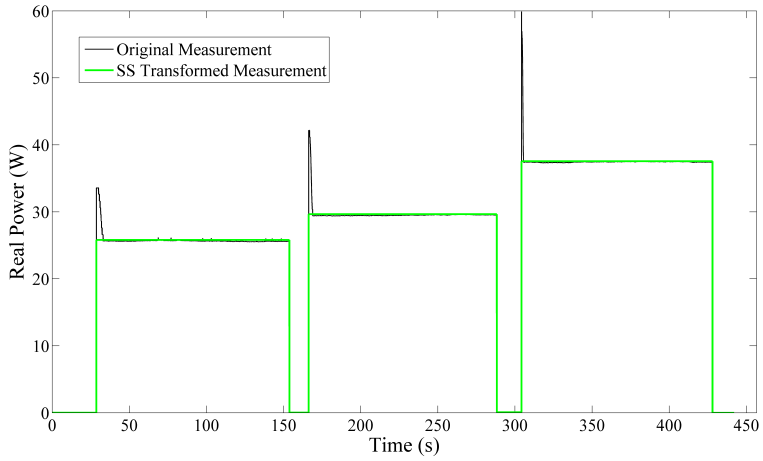

Figure 4.4: Steady-state transformation applied to multi-state fan measurement.

levels. For example, sections of waveform containing ramps may not reach a steady-state before the next appliance state transition occurs, resulting in the calculation of steady-state levels that are considerably higher than should be calculated for the appliances in question. This serves to illustrate the fundamental problem that transients pose to steady-state NILM techniques, namely that many appliances are not steady-state in nature and thus do not fit into such models as might be expected. Steady-states in the power measurements could be detected independently of event edges, by identifying sections of the waveform where power levels do not change dramatically. However, under such schemes, any regions in the measurements containing transients would be excluded from the output, resulting in a large number of appliance activities being ignored. Thus, event edge based approaches offer superior performance and should be adopted ahead of such techniques.

4.3 TLM Implementation

TLM consists of two major phases. The first is the generation of a vector of all feasible combinations of the individual appliance steady-state real power

levels, as extracted from the individual measurements using the steady-state transformation function. The second phase takes the combined measurement real power waveforms after they have been processed by the steady-state transformation function and compares every discrete level against the entries in this vector. In each case the closest match between the two values is taken to indicate the combination of appliances that are operational for the section of combined measurement in question. Phases one and two of the TLM implementation are discussed in Sections 4.3.1 and 4.3.2 respectively.

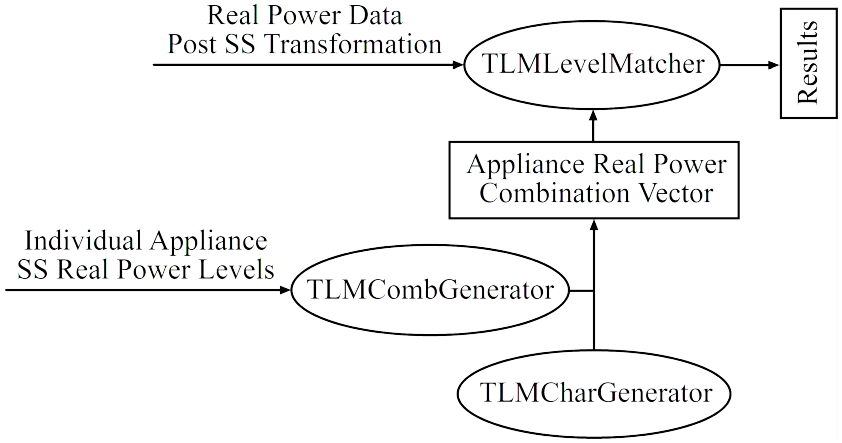

Figure 4.5: TLM implementation overview.

4.3.1 TLM Combination Generator

The *MATLAB* code for this function, `TLMCombGenerator`, may be seen in Appendix C.3. Generating a vector of combinations from the individual appliance power consumption levels is complicated by the presence of multi-state appliances in the laboratory experiment. For example, the fan may not be operating at all three speed settings simultaneously. To handle this problem, multi-state appliances were represented by vectors of exclusive values when generating the combinations, as may be observed in Appendix C.3. The application of this function to the input data produces a vector containing 2048 possible unique appliance state combinations, including the null state where no appliances are operational. This is a substantial num-

ber of combinations, introducing a high potential for error in the matching process, especially given the similar power consumption levels exhibited by many of the appliances.

Given the size of the vector that the `TLMCombGenerator` function produces, it is not trivial to conclude which appliances contribute to the power consumption levels matched to the combined measurement steady-states. Thus some other technique is required for the purposes of providing an identifying label for each entry in the vector. The `TLMCharGenerator` function shown in Appendix C.4 does exactly this, following the same combinatorial process as `TLMCombGenerator`, but using alphabetical characters in place of values.

A unique character is assigned to each appliance state, such that a vector populated with various letter combinations is produced by `TLMCharGenerator`. For example, using a feasible combination of characters taken from the `TLMCharGenerator` function shown in Figure C.4, the characters 'a', 'd' and 'g' may be combined to produce the string 'adg'. As may be seen from the `TLMCombGenerator` function shown in Appendix C.3, this corresponds to a combination of entries 1, 4 and 7 from the 'initMx' vector, or appliance states fan-high, heater-high and microwave-front. Thus, this approach allows for any steady-state level in the appliance combination vector to be easily related back to the appliances from which it was generated.

4.3.2 TLM Level Matcher

The *MATLAB* code for this function, `TLMLevelMatcher`, may be seen in Appendix C.5. Once the steady-state identification approach has been applied to the total power measurement, the unique steady-state levels that are produced must be compared against the 2048 individual appliance combinations generated by the combinator. The `TLMLevelMatcher` function does this by finding the difference between each steady-state level and combination vector entry, where the lowest difference is taken to be the closest

match. The function outputs a sequential list of the individual appliance combination vector row numbers that provided the best matches, which is then used to identify the corresponding appliance combinations by using the alphabetical character vector. A sample of the `compRes` array produced by the TLM process is shown in Table 4.1.

Table 4.1: TLM combination generator output array sample.

Data Point Index	**Measured Steady-State**	**Closest Match**	**Difference**	**Corresponding Appliance Combination**
1	1.48	0.00	1.48	null
495	752.62	753.62	1.00	fnL,ht1,mwF
995	765.45	765.37	0.07	fnH,ht1,mwF
1495	731.57	731.31	0.27	tst
1995	762.18	760.95	1.23	fnM,tst
2502	728.64	728.76	0.11	fnM,ht1,frg,lmp
2996	766.65	767.49	0.84	fnM,ht1,mwI,frg,lmp
3535	725.77	724.92	0.85	fnL,ht1,frg,lmp
11341	0.00	0.00	0.00	null

4.4 CPM Implementation

CPM operates upon the same basic principles as TLM, where steady-state levels extracted from the total power measurement are compared against all possible combinations of individual appliance measurements. The size of the 2048 entry appliance state combination vector that is produced has an adverse effect on the accuracy of the matching process, due to the large number of potential matches that may be available for any measured steady-state power level. In addition to real power measurements, CPM utilises reactive power measurements to obtain an additional layer of information that can be used to refine the appliance state vector, potentially improving disaggregation accuracy.

As discussed in Section 3.2, only two of the appliances included in the laboratory experiment possess reactive power components, namely the fridge and the microwave. The majority of household appliances are purely resistive, containing heating elements and DC motors that typically have power factors of 1 (e.g: the toaster and fan in the laboratory experiment). Between the fridge and the microwave, five reactive power states are available from which to generate the reactive power appliance steady-state vector; null, fridge, microwave-front, microwave-back and microwave-inactive. As the three microwave states are mutually exclusive, this results in CPM generating a reactive power vector with only eight entries, as discussed in Section 4.4.1. Due to the low number of vector levels available to be matched to measured reactive power levels, the accuracy of the reactive power phase of CPM is expected to be significantly better than that of the real power phase.

Once it has been ascertained which reactive power component appliances are present for each steady-state section of the combined measurement, the information can be used to refine the real power appliance combination vector. For example, if the microwave-inactive state is found to be present in the total power measurement based on its reactive power characteristics, then it is reasonable that the real power appliance state combination vector for that steady-state should only contain appliance state combinations that include the microwave-inactive state. In this case, it results in the generation of a real power appliance steady-state combination vector that consists of only 257 entries, which naturally improves the likelihood of an accurate match being made. Even when no reactive power components are present in the total power measurement, an advantage is still offered by this approach, as all of the reactive power appliances may be eliminated from the real power combination vector.

As discussed in Section 3.3.2, the low reactive power level (approximately 30VAR) accompanying the microwave-inactive state often disappears where other appliances are active concurrently. Given that the microwave-inactive state has the lowest power level of the appliance reactive power components

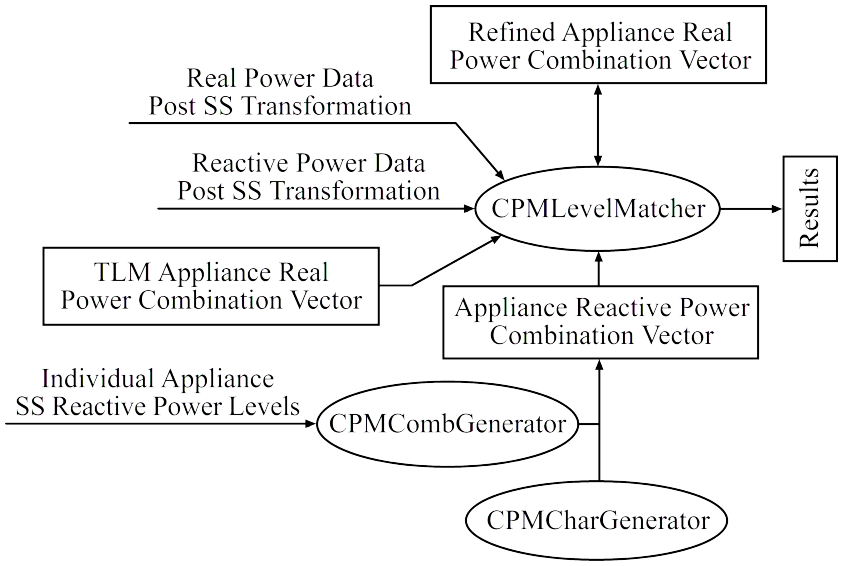

Figure 4.6: CPM implementation overview.

identified in the individual appliance measurements, it is the most commonly affected by inconsistencies in the reactive power measurements. This will introduce error into the CPM process wherever the microwave-inactive state is mistakenly identified as being inactive, as the real power appliance state combination vector that is generated will erroneously omit all appliance combinations containing any of the microwave states.

To address this challenge, two CPM variants were created, one that includes the microwave-inactive state in the reactive power combination vector and another that excludes it. Where it is excluded, the microwave-inactive state must be identified on the basis of its real power component alone. Whilst this addresses the problem of reactive power measurement inconsistencies, it results in a more limited reduction of the real power appliance state combination vector generated for each steady-state, as discussed in Section 5.3.1. Both CPM variants were applied to the combined measurements, with the version that produced the best results being selected as the favoured method, as discussed in Section 5.3.1. The main functions underlying both variants of CPM are presented in the following two sections.

4.4.1 CPM Combination Generator

The *MATLAB* code for this function, `CPMCombGenerator`, may be seen in Appendix C.6. `CPMCombGenerator` operates very similarly to `TLMComb-Generator`, but processes reactive power values instead of real power. As with `TLMCombGenerator`, an alphabetical character vector is generated to assist in interpreting the results. However, for the CPM method, both reactive and real power steady-state appliance combinations must be generated. The real power combinations do not feature all of the appliances, but are refined versions as discussed in the previous section. Thus a series of refined real power combinations, ten in total, must be generated that are employed based upon the reactive power identifications made for each steady-state in the total power measurement. Each of these combinations are created separately using a series of dedicated functions, operating upon a similar basis to both the `CPMCombGenerator` and `TLMCombGenerator` functions, and stored in vectors to be accessed as required by the `CPMLevelMatcher` function presented in Section 4.4.2.

4.4.2 CPM Level Matcher

The *MATLAB* code for this function, `CPMLevelMatcher`, may be seen in Appendix C.7. Again, this function performs in a similar manner to `TLMLevelMatcher`, but factors in reactive power considerations along with real power. As illustrated in Figure 4.1, `CPMLevelMatcher` performs a two stage process for each steady-state detected in the total real power measurement. Firstly the reactive power measurement, which has been processed using the steady-state identification function, is compared to the reactive power appliance combination vector such that the status of the reactive power component appliances may be ascertained. Secondly, a real power appliance combination vector is chosen based upon these findings, and each unique steady-state in the aggregated real power measurement is matched to an entry therein. The process used to perform this real power level matching

is much the same as that employed by the `TLMLevelMatcher` function. However, the `CPMLevelMatcher` output contains both the reactive and real power appliance identifications for each steady-state, allowing for direct observation of the effect of the former on the latter.

4.5 EEC Implementation

The third disaggregation technique under investigation, EEC, involves the comparison of the shapes of the positive and negative edges detected within the total power measurement against corresponding exemplars extracted from the individual appliance measurements. The transient features that characterise the edges of many of the appliance waveforms create distinctive shapes that may be used to distinguish one appliance state transition from another. EEC is performed using real power, as this component is common to all of the appliances. Reactive power could be used to complement this technique. However, the implementation of EEC in this research is intended to assess whether the real power measurement edges possess significantly characteristic shapes to allow the technique to outperform the two steady-state alternatives presented, TLM and CPM. Thus it is undesirable to include additional layers of information, and the sole use of real power measurements is appropriate.

An illustration of the EEC process is given in Figure 4.7. The event detection function, discussed in Section 4.2.1, is employed to identify the edges in the both the individual and combined real power measurements. Once the edges are located, they are captured using the sample extraction function presented in Section 4.5.1. When processing the individual appliance measurements, the microwave sample extractions are handled separately to those performed for the other appliances, using the `EECMWSampleAverager` function discussed in Section 4.5.2. Finally, the individual and combined measurement samples are compared using correlation, where the highest coefficient indicates the best match, as discussed in Section 4.5.3. The event detection

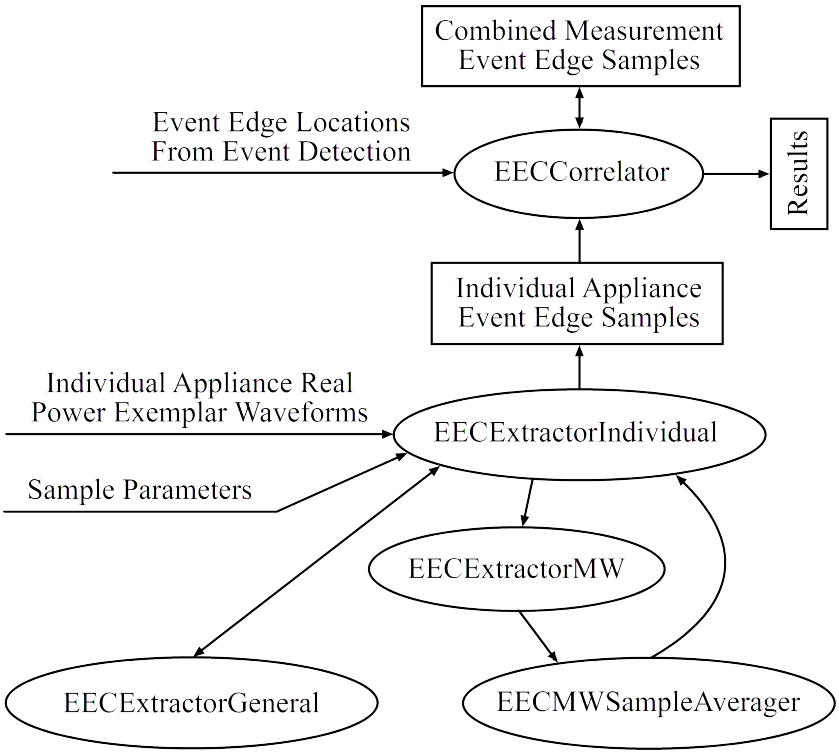

Figure 4.7: EEC implementation overview.

function classifies positive and negative edges separately, which serves to improve the accuracy of EEC. By creating two distinct exemplar libraries, only half the number of edge samples have to be correlated with the total power measurement edges in each case, reducing the potential for mismatches to be made.

4.5.1 EEC Sample Extraction

Two separate functions are employed to extract the samples, one for the individual measurements and another for the combined. The same basic mechanics are used in both cases, with a few distinctions regarding the manner in which the functions are applied to the signals and the results are delivered due to the additional complexity involved in creating the individual appliance edge event sample library. The functions for the indi-

vidual appliance sample extraction process, `EECExtractorIndividual`, `EECExtractorGeneral` and `EECExtractorMW`, are shown in Appendices C.8 through C.10 respectively. All total power measurement sample extraction is performed using a single function, `EECCorrelator`, as discussed in Section 4.5.3. Regardless of the specificities of the individual and combined approaches, both versions of the process take the edge locations in the measurements provided by the event detection function and capture data points to either side of these positions. Four sample parameters must be specified prior to the initiation of the extraction process, as presented in Figure 4.8 and the list below.

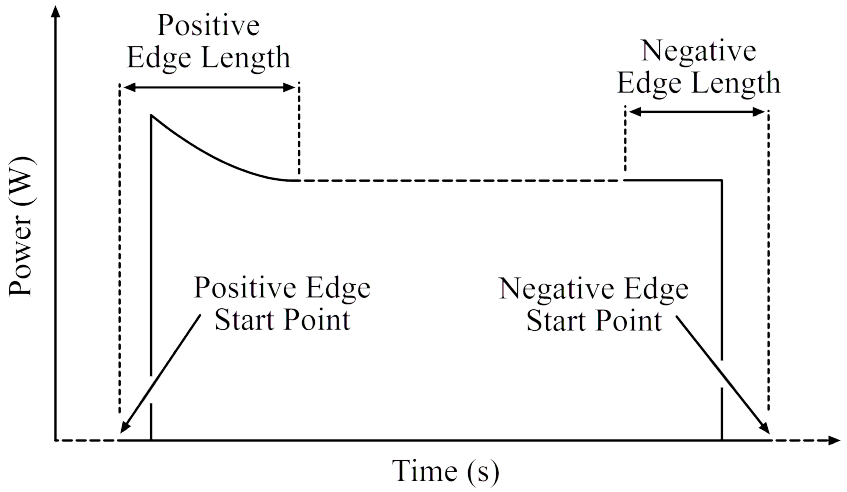

Figure 4.8: Sample parameters presented in the context of an appliance waveform.

Positive Edge Length: The number of data points to be extracted around the position of each positive edge.

Negative Edge Length: The number of data points to be extracted around the position of each negative edge.

Positive Edge Start Point: The number of data points in front of each positive edge position that the extraction should begin.

Negative Edge Start Point: The number of data points in front of each negative edge position that the extraction should begin.

An experiment was conducted to determine the optimum sample parameters, the outcome of which is discussed in Section 5.3.2. When deciding upon potential parameters to trial, the time elapsing between each consecutive appliance state transition events in the combined measurements is a key consideration. On average, 14.5s (725 data points) lie between each event edge, but a minimum of 1.02s (51 data points) is found in Combined Measurement 6, where the fridge is switched off during a microwave-front event. Whilst this is an extreme case, many other examples of intervals below 2s (100 data points) may be found across all nine combined measurements, as shown by the distribution of event gaps presented in Figure 4.9. Four intervals of longer than 25s were excluded from Figure 4.9, being exceptional outliers, with the maximum interval being 156.12s (7806 data points).

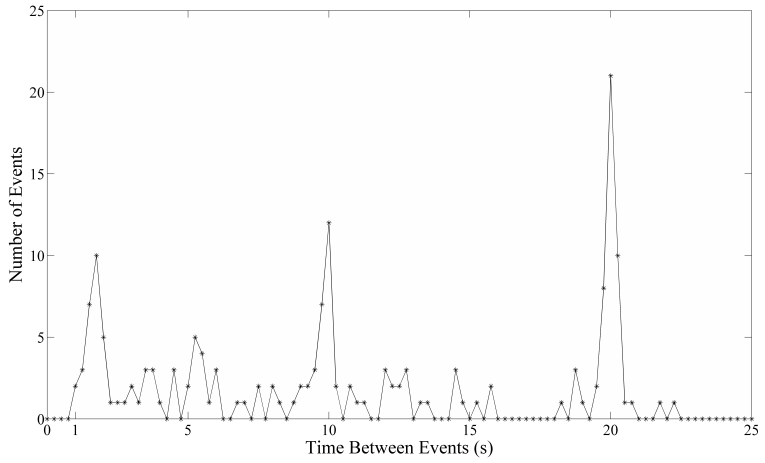

Figure 4.9: Distribution of intervals between events across all combined measurements.

The samples should ideally be long enough to capture the individual characteristic of each edge, but not to the extent that they include other edge events or diminish the weight of any transients within the extracted waveform. However, due to the short intervals that are found between some appliance state transition events, it is unlikely that all of these requirements will be satisfied for each event edge.

The sample start point may impact upon the accuracy of EEC by changing the ratio of high to low amplitude data present in each sample. Where both high and low amplitude data points are captured, they serve to include the magnitude of the power consumption change in the extracted shape. This could potentially enhance the distinctiveness of the samples captured for each appliance state by incorporating the variation that exists between their steady-state real power levels. Alternatively, it could reduce the weight of the event edge transients in the context of the overall samples, making their shapes more homogeneous. These considerations provided the conceptual basis upon which the choices of positive and negative sample start points were made during experimentation to determine the optimal sample parameters.

4.5.2 EEC Microwave Sample Averager

The *MATLAB* code for this function, `EECMWSampleAverager`, may be seen in Appendix C.11. As discussed in Chapter 3, the individual measurements for the microwave medium and low cooking levels contain several magnatron events. None of the events are identical, and thus it is better to construct positive and negative edge exemplars from a series of edge samples than from a single magnatron event. Every time that samples are extracted from the individual appliance measurements, the low, medium and high cooking level microwave waveforms are all processed. This produces a series of positive and negative edge samples for the microwave states which are averaged together into single exemplars using the corresponding data points in each waveform. These averaged microwave waveform edge exemplars are intended to be more widely representative than each of the original state transition edges from which they are compiled.

4.5.3 EEC Correlator

The *MATLAB* code for this function, `EECCorrelator`, may be seen in Appendix C.12. Once the sample extraction process has been completed for each of the individual appliances, `EECCorelator` extracts a sample at every event edge in the total power measurement that is identified by the event detection process, using the same sample parameters. With each extraction, the function correlates the resulting sample against the library of individual appliance event edge samples. As discussed in Section 4.5, positive and negative edges are processed separately to improve accuracy. The appliance sample that scores the highest correlation coefficient in each case is taken to be the best match, and hence indicates which appliance was responsible for the event in the total power measurement. A sample of the `combP` array produced by the concluded EEC process is shown in Table 4.2.

Table 4.2: EEC correlator raw output sample, showing data point indices, edge directions and appliance correlation values.

Appliance	**+495**	**+995**	**-1495**	**+1995**	**-2502**	**+2996**
Fan-High	0.8357	0.5967	0.9444	0.7251	0.9160	0.9561
Fan-Medium	0.9610	0.8015	0.9695	0.9925	0.9358	0.8234
Fan-Low	0.8617	0.9527	0.9704	0.8748	0.9522	0.5833
Heater-High	0.1210	-0.0463	0.7992	0.1172	0.7812	0.1030
Heater-Medium	0.1241	-0.0436	0.8101	0.1201	0.7913	0.1059
Heater-Low	0.1110	-0.0452	0.9664	0.1115	0.9345	0.0880
MW-Front-1	-0.8946	-0.8301	0.0000	-0.9707	0.0000	-0.6922
MW-Front-2	0.0947	0.4597	0.3441	0.1812	0.4206	-0.1595
MW-Inactive	0.4166	0.3853	0.9184	0.3465	0.8921	0.4965
Fridge	0.0295	0.1104	0.9656	-0.0045	0.9356	0.0387
Kettle	0.1829	0.0091	0.6872	0.1753	0.6755	0.1576
Lamp	-0.0580	0.0775	0.9229	-0.0741	0.8966	-0.0361
Snackwich	0.1104	-0.0561	0.9670	0.1079	0.9387	0.0934
Toaster	0.9896	0.8348	0.9648	0.9507	0.9309	0.8757

Each correlation is performed using the *MATLAB* function `corr`, which tests the degree of association between the two sample vectors being compared.

A coefficient of 1 indicates perfect correlation, implying that the two edge samples are identical. A coefficient of -1 indicates the same, but with a phase shift of 180°. As the coefficient tends towards 0, so the two sample vectors being compared exhibit decreasing levels of similarity. As a default, the `corr` function within *MATLAB* performs correlations using Pearson's linear correlation coefficient, with Kendall's Tau and Spearman's Rho available as alternative options [35].

Table 4.3: Accuracies of correlation coefficient types for sample parameter combinations 'Par1' and 'Par2'.

	Pears. (Par1) (%)	**Kendl. (Par1)** (%)	**Spear. (Par1** (%))	**Pears. (Par2)** (%)	**Kendl. (Par2)** (%)	**Spear. (Par2)** (%)
Comb. Meas. 1	75	13	13	75	13	13
Comb. Meas. 2	59	53	59	59	59	59
Comb. Meas. 3	55	52	52	38	45	55
Comb. Meas. 4	64	14	7	50	21	7
Comb. Meas. 5	38	0	0	38	0	0
Comb. Meas. 6	58	39	39	52	35	45
Comb. Meas. 7	7	14	14	14	14	14
Comb. Meas. 8	27	0	0	9	0	0
Comb. Meas. 9	43	37	43	27	33	40

The Pearson, Kendall and Spearman coefficients are all commonly employed in performing correlations, their selection depending on the nature of the data being compared. However, Pearson's coefficient is the most widely used as it is intended for application to normally distributed variables, whereas the other two approaches are intended for non-normally distributed data [36]. To assess which would be the most appropriate for use with the edge samples, EEC disaggregation passes were conducted using all three correlation approaches. Two well performing sets of sample parameters were used, both featuring substantial positive and negative sample lengths. However, one set contained a mixture of high and low amplitude data points, whereas the other contained only high amplitude data. This was done in order to evaluate how each of the correlation coefficient types performed when applied to the two

different types of sample shapes.

Although not 100% consistently returning the best accuracy for each edge sample correlation, Pearson's coefficient produced the most accurate results overall, as shown in Table 4.3, and hence was selected ahead of the Kendall and Spearman coefficients.

This concludes the presentation and discussion of the TLM, CPM and EEC implementations. The results obtained from applying the disaggregation techniques to the measurements are presented in Chapter 5.

Chapter 5

RESULTS AND OBSERVATIONS

The results obtained from applying TLM, CPM and EEC to the combined measurements are presented. The methods used to score the accuracy of each technique are discussed, along with the CPM variant and EEC sample parameters selected for inclusion in the final results and the validity of comparisons made between each technique. An overview of the results is presented, followed by a series of key observations drawn out of the detailed results analysis and discussion contained in Appendix G.

5.1 Results Process Overview

Before the final set of results may be produced for the comparative evaluation of TLM, CPM and EEC, as presented in Section 5.5, a series of steps must be followed; disaggregation scoring methods must be developed, the best CPM

variant selected and the optimal EEC sample parameters determined. This process is illustrated in Figure 5.1.

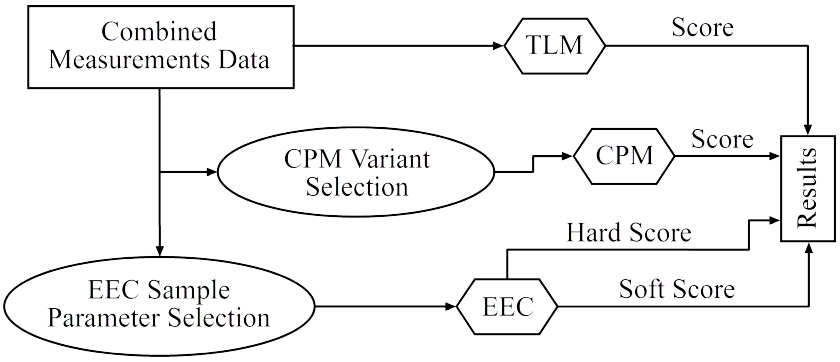

Figure 5.1: Overview of the remaining steps required to produce the final results.

The scoring methods used for evaluating the accuracy of each disaggregation technique need to be designed such that the final results produced by TLM, CPM and EEC may be comparatively evaluated, as discussed in Section 5.2. Due to the reactive power measurement inconsistencies discussed in Section 3.3, two variants of CPM were developed, as presented in Section 4.4. The variant that includes the microwave-inactive state offers the best disaggregation accuracy, and is thus selected for the production of the final results, as discussed in Section 5.3.1. Furthermore, the optimal sample parameters for EEC, as determined through experimentation, are presented in Section 5.3.2.

5.2 Scoring Methods

The scoring method used to interpret the EEC results must necessarily be different from that used for TLM and CPM, due to the fundamental differences between transient and steady-state appliance signatures. Accordingly, scoring methods were developed such that the disaggregation performances of each of the three techniques may be directly compared, as discussed for TLM and CPM in Section 5.2.1 and for EEC in Section 5.2.2.

5.2.1 TLM and CPM Scoring

TLM and CPM are scored identically, utilising a binary approach. Under this scheme, a score of either '1' or '0' is allocated to the disaggregation outcome for each steady-state in the total power measurements, where '1' represents a perfectly correct appliance state combination identification and '0' an imperfect result. Should an outcome be only partially accurate, perhaps correctly identifying one or more appliances in a particular steady-state but incorrectly identifying the remainder, the score is still taken to be '0'. Whilst this is a relatively strict approach to scoring that does not recognise partially correct outcomes, the mechanism underlying TLM and CPM identifications necessitates that only fully accurate results be rewarded.

Given the combinatorial matching technique used to perform TLM and CPM identifications, any partially correct solution may be considered to be no better than a completely erroneous one. When the wrong entry is selected from the vector of appliance state combinations, any similarity between the contents of the entry and actual appliance operations is incidental, as both TLM and CPM consider only the total power 'best fit' for each steady-state, with no capacity to discern between the actual constituent elements of the measurement. Whilst the two layers of combinatorial matching used for CPM could be considered grounds for the recognition of partial identifications, the reactive power stage of the process only serves to alter the size and contents of the second stage of matching, and thus partially correctly appliance state identifications are still largely incidental.

EEC is evaluated using a binary scoring approach, as discussed in Section 5.2.2. Thus the use of binary scoring for CPM and TLM allows for the direct comparison of all three disaggregation techniques, given the implementation of the additional compatibility measures presented in Section 5.4.

5.2.2 EEC Scoring

Two binary scoring methods are used for assessing the performance of all EEC appliance state transition identifications, namely a 'hard' and 'soft' score. As EEC considers each event edge in the total power measurements to be associated with a single appliance, a partially correct solution may not be returned in the same manner as might be for TLM and CPM. However, given that some of the appliances possess multiple operating states, it is possible that an appliance be correctly identified, but the incorrect state attributed to it. For example, activation of the fan may be accurately detected, but the chosen motor speed may not be correctly recognised. The hard scoring system only awards a '1' where both the appliance and its operating state are correctly identified, whereas the soft approach will assign a '1' if the appliance is identified, regardless of whether its operational state is accurately recognised or not. Thus the use of the soft scoring system only impacts upon the evaluation of multi-state appliances. For both approaches, a '0' is assigned wherever the criteria for a successful identification are not met.

Whilst a partially correct identification score is not valid for TLM and CPM, as discussed in Section 5.2.1, it is justifiable for EEC. Although each state of a multi-state appliance may have different magnitudes, the electrical components and physical activities of the appliance produce event edge transients that have similar characteristics, regardless of the state. For example, each of the three fan speed setting states have a similar transient peak that accompanies their waveform positive edges, as shown in Section 3.2.6. Thus, it is valid to credit EEC with a partially correct identification in cases where an appliance is correctly identified, but assigned the incorrect operational state. This is due the fact that the characteristic shape of the appliance waveform event edge in question has been used to identify the appliance, regardless of the assignment of the incorrect operational state, and thus the partial identification is not a random outcome.

The hard scoring system is of primary interest here, being more relevant to

the research being conducted than the soft approach, as the identification of precise appliance multi states is more consistent with the initial aims of the experiment. Furthermore, as a hard approach is utilised for scoring the TLM and CPM results, it is appropriate that this be the main scoring method for EEC, ensuring that any comparative evaluations of the scores remains valid. However, the soft scoring method has been retained as an additional metric for EEC, as it does offer an indication of how the technique might perform in situations where it is not required to distinguish between appliance states.

5.3 Variants and Parameters Selected

As discussed in Section 4.4, two variants of CPM were developed, one of which must be selected to produce the final results. Additionally, the optimal parameters for extracted samples for EEC were ascertained via experimentation, as per Section 4.5.1. Both of these processes, and their outcomes, are presented below.

5.3.1 CPM Variant

Comparative results for the two CPM variants are shown in Table 5.1. Where incorrect identifications of the reactive power component of the microwave-inactive state occur, the variant that includes the microwave-inactive state in the reactive power matching process will introduce error into the results. However, it may be seen that this variant still either equals or outperforms the alternative in all cases, except for Combined Measurement 9, and may be considered the best performer overall, as shown in the bottom five entries of Table 5.1. Thus this CPM variant is selected for the production of the final results. This outcome is expected, given that the inclusion of the additional state in the reactive power matching process allows for greater refinement of the real power appliance combination vector; where the selected variant

produces vectors with 257 entries, whilst the alternative variant produces vectors with 513 entries, approximately double the size.

Table 5.1: Accuracy results for CPM variants.

	Included (%)	**Excluded (%)**
Comb. Meas. 1	50	38
Comb. Meas. 2	47	35
Comb. Meas. 3	38	28
Comb. Meas. 4	71	71
Comb. Meas. 5	100	100
Comb. Meas. 6	32	26
Comb. Meas. 7	64	43
Comb. Meas. 8	9	9
Comb. Meas. 9	32	42
Best of Nine	**8/9**	**4/9**
UWA	**49**	**43**
UWA excl 5 & 7	**40**	**36**
EWA	**46**	**41**
EWA excl 5 & 7	**40**	**36**

Table 5.2 shows the results for Combined Measurement 1 using CPM with the microwave-inactive reactive power component excluded from the matching process. It may be seen that the toaster and fan-high combination found between 60s and 71s was only partially correctly identified, where the toaster was mistaken for a combination of the microwave-inactive and snackwich operational states. However, under the chosen CPM variant this was not the case, as the microwave-inactive state was eliminated from the vector of possible real power appliance state combinations through the reactive power matching process.

As shown in Table 5.2, the combined snackwich and microwave-inactive real power consumption steady-state level, found between 50s and 60s, is 728W and the toaster consumes 731W alone. Such similar steady-state levels may easily become confused, as discussed in Section 5.6. This makes any additional information that may be used to refine the appliance state combination vector useful to the disaggregation process and validates the choice of CPM

Table 5.2: CPM accuracy results for Combined Measurement 1 with microwave-inactive state excluded.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
10	-0.47	0.00	null	752.62	753.93	fnL-mwI-snw	tst	0
20	0.00	0.00	null	765.45	765.68	fnH-mwI-snw	tst,fnL	0
30	0.00	0.00	null	731.57	731.31	tst	tst	1
40	0.00	0.00	null	762.18	760.95	fnM-tst	tst,fnM	1
50	0.00	0.00	null	728.64	728.12	mwI-snw	tst	0
60	0.00	0.00	null	766.65	765.68	fnH-mwI-snw	tst,fnH	0
71	0.04	0.00	null	725.77	726.95	fnH-snw	tst	0
227	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**38**

variant, despite the error that it introduces where the microwave-inactive reactive power steady-state is incorrectly identified.

5.3.2 EEC Sample Parameters

Table 5.3 shows comparative results for a range of sample parameters used with the EEC method. In each case, the length of the positive and negative samples and their starting positions relative to the event edges within the total power measurements have been varied in order to find the best performing parameters, as discussed in Section 4.5.1. The following parameters may be seen to return the best performance; positive edge length of 250 data points, negative edge length of 200 data points and no time shift on either the positive or negative event edge extraction points. An illustration of the optimal sample parameters is presented in Figure 5.2.

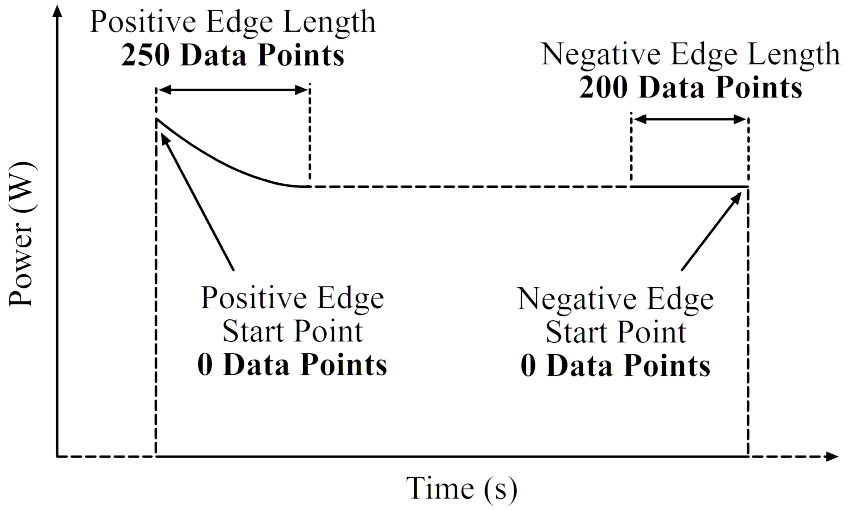

Figure 5.2: Optimal sample parameters presented in the context of an appliance waveform.

As shown in Table 5.3, the chosen sample parameters are marginally outperformed by other combinations of parameters for combined measurements 7 and 9. However, for combined measurement 7 all sample types performed poorly, and thus the marginally poorer performance of the chosen parameters

Table 5.3: EEC accuracy results for various sample parameters, in format +length/-length/+start/-start.

	075/075/ 00/00 (%)	**100/100/ -15/15** (%)	**150/150/ -10/10** (%)	**250/200/ 00/00** (%)	**250/200/ -10/10** (%)	**250/200/ -50/50** (%)	**300/250/ -10/10** (%)
Comb. Meas. 1	63	63	63	75	75	75	75
Comb. Meas. 2	29	47	53	59	59	53	59
Comb. Meas. 3	55	52	48	55	45	48	38
Comb. Meas. 4	50	36	43	64	50	43	50
Comb. Meas. 5	19	19	31	38	38	38	38
Comb. Meas. 6	39	48	52	58	55	45	52
Comb. Meas. 7	0	14	14	7	14	14	14
Comb. Meas. 8	27	9	9	27	9	9	9
Comb. Meas. 9	33	37	47	43	27	43	27
UWA	**35**	**36**	**40**	**47**	**41**	**41**	**40**
EWA	**36**	**38**	**42**	**47**	**41**	**42**	**39**

does not invalidate the choice made. In the case of combined measurement 9, the best performing sample parameters only marginally outperformed the chosen set of parameters and thus this result also does not invalidate the choice made. The chosen sample parameters either outperformed the alternative options or shared the same score in every other case, returning the highest average score across all of the combined measurements. Whilst these chosen parameters could well provide good results when applied to other research, further experimentation with different data sets would have to be conducted in order to consider them other than experiment specific.

This result is not expected, given that there is no 'low' data included in the samples, due to the lack of time shifting around the event edges in the combined measurements. The inclusion of 'low' data points might be expected to make the samples more easily distinguishable from one another, as discussed in Section 4.5.1. However, it may be seen from the results in Table 5.3 that this is not the case, and that performance is enhanced by the inclusion of 'high' data points only.

5.4 Technique Comparison Considerations

By virtue of their similarity, and the scoring approaches used, the accuracies of TLM and CPM may be directly compared. However, EEC performs appliance state identifications using a different characteristic than TLM and CPM, and thus its results may not be directly compared to those of the other two techniques without adjustments being made. TLM and CPM produce a new combination of appliance states for each steady-state found in the combined measurements, with no reference to previous identifications. In contrast, EEC identifies the appliance responsible for each transition between steady-states, without identifying the combination of appliance states active after the edge has passed. Thus the key to allowing for inter-technique comparisons to be made lies in the assumption that the sequence of EEC identifications that has been made in the lead up to each new event edge

in the total power measurements is 100% accurate. This means that each new event processed by EEC may be considered to be a unique event and is judged upon its individual merit, as is the case with both TLM and CPM.

For example, if a fridge negative edge is identified by EEC, then what is truly being compared against the corresponding TLM and CPM identifications is the steady-state condition that exists after the fridge has been deactivated. Whereas TLM and CPM will produce a combination of devices that matches this steady-state, the equivalent EEC result is considered to be the known combination of appliances from the ground truth data, minus the fridge. If the edge is not correctly identified, then the incorrect combination of appliances will be considered to be operational after that negative edge, and is thus equivalent to a poor TLM or CPM match for the same section of the measurement. It is important to note that EEC is implemented on the assumption that no appliances are active prior to the first detected event, which would be a positive edge. To directly produce combinations of appliance states using EEC, memory of the edge identifications could be held throughout the course of a disaggregation. Whilst this could be advantageous, where possible sample matches could be refined using logic, any incorrect identifications made under such a scheme would permeate throughout the rest of the process, introducing further error into the results, as discussed in Section 6.3. Accordingly, EEC is evaluated on a case-by-case basis that allows for each new event edge in the total power measurements to be treated as a unique disaggregation, enabling direct comparison of its performance with that of TLM and CPM.

All null states found within the combined measurements are correctly identified by TLM and CPM. This may not have been the case had the real power noise levels in the laboratory experiment not remained below the 15W event detection threshold throughout all nine combined measurements. The 100% accurate identification of null states serves to bolster the reported accuracy of the TLM and CPM methods, especially where multiple null states are found in one combined measurement. Given that EEC does not directly recognise

nulls, only the negative edges that precede them, the null states found at the start of each of the combined measurements are discarded. This means that any disaggregation scores occurring prior to the first event edge in the total power measurement are excluded from the TLM and CPM results. All other null states within the combined measurements were retained, as shown in Figure 5.3 and its accompanying Table 5.4.

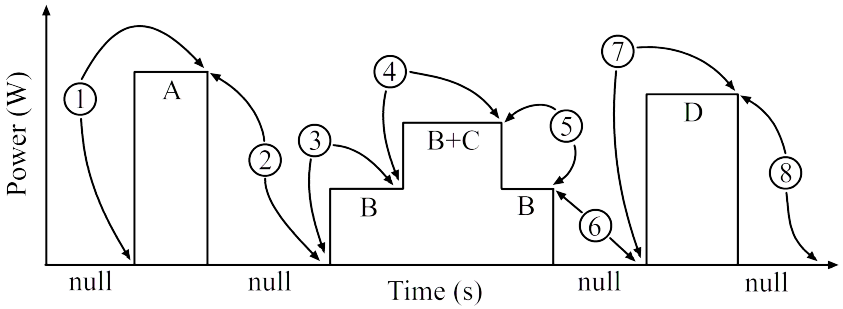

Figure 5.3: Example of interaction between steady-states, event edges and null states, to accompany Table 5.4, featuring fictional appliances A through D.

Table 5.4: Method of inclusion of steady-states and event edges in results, to accompany figure 5.3.

Section	**TLM and CPM**	**EEC Positive Edge**	**EEC Negative Edge**
1	A is ON	A turned ON	–
2	Null state	–	A turned OFF
3	B is ON	B turned ON	–
4	B and C are ON	C turned ON	–
5	B is ON	–	C turned OFF
6	Null state	–	B turned OFF
7	D is ON	D turned ON	–
8	Null state	–	D turned OFF

In many of the combined measurements, certain appliance state combinations are repeated, including null states. These repeated combinations are

retained, regardless of the extent to which they may reoccur. This serves to simulate real-world conditions, where repeated appliance state combinations are to be expected, and to test the consistency of each technique being applied. Whilst there is significant merit in the idea of testing every possible combination of appliances included in the laboratory experiment, this is a time consuming undertaking that would likely be prone to considerable human error, given how extensive such an experiment would be. Accordingly, the ground truth data was developed to simulate nine feasible sequences of appliance state combinations, including repeats. An alternative approach to the measurements methodology that could provide more extensive data, including all possible appliance state combinations, is proposed in Section 6.2.

5.5 Overview of Results

This section contains a summarised analysis of the results obtained from the application of TLM, CPM and EEC to the combined measurements. A more detailed discussion of the performance of each technique for each combined measurement may be found in Appendix G.

Table 5.5 shows that combined measurements 5 and 7 feature a combination of excellent results for TLM and CPM, and poor accuracy for EEC. This is particularly so in the case of Combined Measurement 5, where TLM and CPM both achieved perfect results. These may be considered as outlier performances that are not representative of the results in general, given the reported TLM and CPM accuracies for the other combined measurements. As these two combined measurements skew the results significantly, the overall accuracies should be considered both with and without their inclusion. The results shown in Table 5.5 may be interpreted using a number of approaches, each of which applies a different averaging method to the individual disaggregation instances that constitute the combined measurements, as discussed below.

Table 5.5: Combined measurement accuracy results for TLM, CPM and EEC.

	TLM (%)	**CPM (%)**	**EEC Soft (%)**	**EEC Hard (%)**
Comb. Meas. 1	38	50	100	75
Comb. Meas. 2	12	47	59	59
Comb. Meas. 3	17	38	66	55
Comb. Meas. 4	57	71	79	64
Comb. Meas. 5	100	100	56	38
Comb. Meas. 6	16	32	61	55
Comb. Meas. 7	43	64	14	7
Comb. Meas. 8	9	9	27	27
Comb. Meas. 9	23	32	53	43
Best of Nine	**1/9**	**3/9**	**7/9**	**6/9**
Best of Seven	**0/7**	**1/7**	**6/7**	**6/7**
UWA	**35**	**49**	**57**	**47**
UWA excl 5 & 7	**24**	**40**	**64**	**54**
EWA	**31**	**46**	**57**	**47**
EWA excl 5 & 7	**24**	**40**	**61**	**52**

Best of Nine: This approach assigns a point to the disaggregation technique which performs the most accurately for each combined measurement. In the case of a tie, both techniques are granted a full score. According to this approach, EEC is the best performer, followed by CPM and then TLM. With combined measurements 5 and 7 excluded, EEC leads even more comprehensively, followed by CPM and TLM.

Unequal Weighting Average (UWA): This approach takes the average scores for each of the combined measurements and finds the overall average of those. Under this scheme, each combined measurement carries equal weighting regardless of the number of identifications that it contains. Thus each individual disaggregation is not equally weighted, with greater weighting given to the accuracies of combined measurements containing fewer appliance state change events. If the UWA is adopted, CPM emerges as the best of the techniques, marginally ahead of EEC, and followed by TLM. However, with

measurements 5 and 7 excluded, EEC takes the lead, followed by CPM and TLM.

Equal Weighting Average (EWA): This is the average of all of the individual disaggregation attempts made across all of the combined measurements, without any consideration being given to the particular combined measurement within which each performance is contained. This results in each individual disaggregation instance being weighted equally across the entire set of results. Under this approach, EEC marginally outperforms CPM in terms of accuracy, followed by TLM. However, with combined measurements 5 and 7 excluded, EEC offers the best accuracy, followed by CPM and TLM.

The most representative average is the EWA with 5 and 7 excluded, which provides the most generalised probability that a correct outcome will be gained for any single identification being performed. Thus two observations may be drawn from these results; the first being that EEC is the best performing disaggregation technique and the second that CPM provides an improvement on the accuracy of TLM. This may be extended to say that the comparison of transient event edges in the total power measurements provides a better disaggregation approach than the comparison of steady-state power levels, and that the consideration of both real and reactive power data in the context of steady-state power disaggregation is superior to the exclusive use of real power data.

Analysis of the overall results leads to the auxiliary question; "Have any of the techniques been useful in disaggregating the total power measurements?" Whilst the results obtained do comparatively evaluate the three disaggregation techniques, satisfying the primary aim of the research, all of the techniques would need to be developed and refined further before being implemented in any real-world NILM system. A high identification accuracy is required for such systems, where instances of poor performance will negatively affect user trust levels and hence reduce their willingness to engage

with the system. However, as this study is purely comparative, the absolute magnitude of the accuracies are of less interest than their relative values.

Of the three disaggregation techniques, EEC is the most promising for future research and development. This conclusion is based upon the overall accuracy results and the observations presented in Section 5.6. The addition of appliances into the circuit under test will adversely affect the accuracies of TLM and CPM, due to the high number of new combinations that must be generated for comparison to each measured steady-state. Whilst the performance of EEC will also be affected by the inclusion of further appliances, the effect is relatively marginal as the technique does not utilise appliance state combinations. This consideration serves to promote EEC further as the most promising of the disaggregation techniques for further research and development.

5.6 Observations Drawn from Results

The following sections contains observations drawn from the detailed results discussion conducted in Appendix G.

Similarity of Power Consumption Levels

Wherever appliance states, or appliance state combinations, feature similar power consumption levels, there is potential for them to be mistaken for one another; especially given the variability observed in the total power measurements. Many examples of the error introduced by this phenomenon, perhaps the greatest impediment to the performance of the steady-state disaggregation techniques, are evident in the results and discussion contained in Appendices F and G respectively. When an incorrect match is made, the techniques themselves are not truly at fault, as they have taken the measured steady-state level and successfully found the mathematically closest match.

Rather, it is the similarity in power levels of alternative matches in the appliance state combination vector that leads to the selection of a steady-state combination that differs from the ground truth data and is thus deemed to be incorrect.

Consequently, the lower the levels of similarity found between the various appliance state combinations contained within the circuit being disaggregated, the more accurate the results are likely to be. This means that the performance of these techniques is heavily dependent upon the power consumption characteristics of the particular appliances included within the electrical circuit being measured. Under favourable conditions, where the appliance-state combination vector values are widely and evenly spaced, good disaggregation accuracies could be realised using TLM and CPM. However, real-world scenarios are unlikely to provide such a fortunate state of affairs, especially where larger number of appliances are involved. This has an adverse affect on the potential of either TLM or CPM to provide an accurate real-world disaggregation solution.

Appliance Combination Vector Favourable Matching Regions

The accuracy of both TLM and CPM is linked to the magnitude of each measured steady-state real power level, where the higher the magnitude of the measured level, the lower the likelihood of an accurate identification being made. The results contained in Appendix F bear out the observation that lower steady-state real power levels are easier to correctly match than higher values, showing that a favourable region for making identifications exists towards the lower end of the vector of appliance state combinations. Any particular combination of appliances within a circuit being measured will produce ranges within the minimum and maximum possible combined values (from 0W to 5992W for the laboratory experiment) that feature more potential appliance state combinations than other ranges do. This means that a measured level that falls into a sparsely 'populated' range is more likely to

be accurately identified than one that falls into a densely 'populated' range, given that the probability of error is increased when a measured level has many potential matches of a similar value.

Favourable regions for matching should be expected to be found around the lower and upper ends of the possible range of combined values. However, it is unlikely that higher value favourable regions be reached as often as the low valued ones, given that this requires that almost all of the appliances be operational at one time. Thus, practically speaking, the main favourable matching region of consequence would be expected to be located with the lower end of the range, with accuracy worsening as measured power values increase. Certainly, this has been found to be the case with the laboratory experiment, a practical example of this effect being given in Appendix G.5. Both the TLM and CPM results are influenced by this consideration, as the techniques utilise the same approach to real power matching in order to make identifications. CPM improves on TLM primarily by reducing the size of the vector of combinations to be matched. However, implicit in this refinement is the expansion of the real power favourable matching region such that the matching of higher measured levels is improved, as evidenced in the results.

Table 5.6 contains basic statistics for each of the combined measurements for TLM and CPM. It may be seen that in most cases where there is an improvement in the count of correctly identified entries from TLM to CPM, a higher real power maximum level is accurately matched by CPM than TLM. Whilst this is a function of the reduced appliance state combination vector, it may be noted that the median value is exceeded by the highest matched CPM real power level in all but one case. However, the TLM maximum matched level is either less than or approximately equal to the median for eight out of nine of the combined measurements. This provides an indication of how the favourable region for accurate matching has been expanded by CPM, allowing for higher measured real power levels to be correctly identified than possible with TLM.

Table 5.6: Collated TLM and CPM statistics for all combined measurements.

	Count	**Mean (W)**	**Maximum (W)**	**Median (W)**	**TLM Successes**	**TLM Maximum Match (W)**	**CPM Successes**	**CPM Maximum Match (W)**
Comb. Meas. 1	8	747.56	766.65	742.10	3	762.18	4	766.65
Comb. Meas. 2	17	733.10	1748.70	766.41	2	39.84	8	1024.72
Comb. Meas. 3	29	761.70	1842.69	766.46	5	769.66	11	1056.27
Comb. Meas. 4	14	430.66	778.23	397.05	8	745.94	10	745.94
Comb. Meas. 5	16	437.46	617.06	565.41	16	617.06	16	617.06
Comb. Meas. 6	31	922.78	2547.29	945.84	5	320.23	10	1072.39
Comb. Meas. 7	14	917.47	2085.08	787.62	6	790.33	9	1350.75
Comb. Meas. 8	11	1789.33	2212.47	2147.15	1	2003.83	1	2003.83
Comb. Meas. 9	31	1221.77	2567.62	1293.33	7	936.25	10	1527.03

Variability in Measured Power Levels

A combination of noise and hidden or variable states not encapsulated in the individual measurements serves to introduce discrepancies between the expected and measured power levels for each appliance state, or appliance state combination. Whilst the noise level in the real power measurements never exceeds 15W, there are still instances where an incorrect identification may be directly attributed to noise. An example of such a case may be seen between 210s and 230s in Appendix E.2 for the reactive power matching stage of CPM, and the same issue affects real power matching for both TLM and CPM. Furthermore, an appliance may unexpectedly switch into a combination of hidden states that were not detected in the individual measurements, introducing error into the disaggregation process. This state of affairs could occur because the states in question are a rare event that did not occur during individual measurement, or because the measured appliance states contain some degree of variability that either went undetected or could not be catered for adequately by the disaggregation technique in question. The constant downward slope of snackwich real power consumption is an example of such variability, where the period for which the snackwich is active affects the steady-state power consumption level that is measured for it, as discussed in Section 3.2.5. These inconsistencies between the measured and expected steady-state levels make it difficult for TLM and CPM to make correct identifications, especially when the measured power level is not in a favourable matching region, as discussed in Section 5.6.

Another example of such steady-state variability may be found in Combined Measurement 1, where the initial ramp following activation of the toaster is interrupted by a fan-low state at 20s. The waveform section of interest is shown in Figure 5.4, and the full combined measurement may be seen in Section 3.3.1. As the toaster waveform did not have time to settle down to the true steady-state, the calculated value for this section of the total power measurement is significantly higher than would be expected when referring to the ground truth data. As a consequence of this, that particular

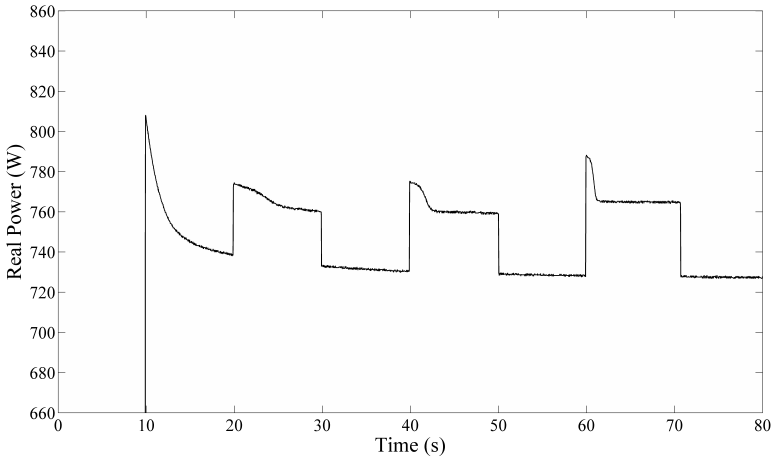

Figure 5.4: Magnified section of Combined Measurement 1 waveform.

appliance state was incorrectly identified by both CPM and TLM. Figure 5.4 also includes fan-medium and fan-high waveforms, starting at 40s and 60s respectively, which can be seen to be located closer to the steady-state section of the toaster waveform, and are both correctly identified by CPM. Due to the variation in time lengths between peaks, the mean calculation method used for the steady-state transformation of the measurements provides the best approach to the identification of steady-states. If steady-state detection was used instead of this method, no steady-state would have been found for this section and hence no appliance identification could have been performed at all using TLM and CPM. Thus it may be seen that the time between events also has an effect on the accuracy of steady-state power disaggregation techniques, given that event edges often contain ramps, slopes and transient peaks that take time to settle down to steady-state levels.

Variability in measured reactive power levels, as discussed throughout Section 3.3, adversely affects the accuracy of CPM for combined measurements 2, 3, 6, 8 and 9. The error that such variabilities introduce into the reactive power matching stage of CPM is carried through to the reduced size real power appliance state combination vectors produced by the technique. Matching of the microwave-inactive reactive power state was the most af-

fected by these inconsistencies, due to the low reactive power level that it consumes, prompting the investigation of CPM two variants aimed at improving the accuracy of microwave-inactive state identifications, as discussed in Section 4.4.

In real-world implementations, other external factors may introduce additional variability into the power measurement, further affecting steady-state power consumption disaggregation techniques such as TLM and CPM. For example, if a heavy object is placed in the microwave it may draw more power than with a lighter load, due to the higher power level required to rotate the internal tray. Whilst the microwave was operated without a cooking load throughout this experiment, such variations in measured power consumption make accurate steady-state power level matching even more problematic in real-world implementations, providing additional motivation for the further development of EEC ahead of TLM and CPM.

Distinctiveness of Event Edges

As discussed in Appendix G, microwave event edges are identified with a high level of consistency throughout the EEC results. EEC outperforms TLM and CPM for combined measurements 2, 3, 6 and 9, all of which feature microwave operations. Based upon the principles of the technique, it might be reasoned that the distinctiveness of the microwave edge events is responsible for the high accuracy exhibited in these cases. The toaster also features a distinctive positive edge in the context of the appliances included in the laboratory experiment. Whilst it is identified with 100% accuracy in Combined Measurement 1, it is only identified with 50% accuracy in Combined Measurement 7. This inconsistent accuracy may be ascribed to the fact that toaster waveform positive edge is not as distinctive as the microwave positive event edges, making it more problematic for EEC to identify toaster activation events consistently.

In Combined Measurement 5, where the appliances utilised do not possess

particularly characteristic edges, EEC only offers 38% accuracy, again pointing to the influence of this factor on the performance of the technique. Thus, drawing upon the EEC discussions conducted throughout Appendix G and the points mentioned above, it may be argued that the shape of appliance edge events exert a major influence on EEC. This in turn means that EEC will work better in cases where the appliances contained in a circuit feature high variation in the shape of their event edges, leaving its performance subject to the characteristics of the appliances under test. This reliance on a particular characteristic of each event provides a strong positive that may be associated with EEC and similar transient techniques. It implies that if the method of comparison or extraction of the shape characteristic is refined further, such that the distinctiveness of each extracted edge sample is emphasised, then the technique will return improved performances.

Variability may be found between different event edge samples generated by the same appliance. For example, the microwave front exemplar samples had to be compiled into an average sample in order to compensate for the disparities between magnatron events occurring within the same measurement period. This variability, combined with the effects of noise, adversely affects the accuracy of EEC. The correlation figures presented in Appendix F illustrate the variation found between individual edge samples produced by the same appliance, which might be expected to exhibit perfect correlation.

Similar Correlation Values

Two problems may be observed within the correlation values produced by EEC, as contained in Appendix F. Firstly, there are many instances where low correlation figures are found for all appliances being compared against a particular event edge. In such situations, where a correct identification is made, it is not done so with any measure of confidence. Rather, the correlation figures indicate that none of the individual samples match the captured edge, but that the chosen one is the best of the mismatches. This

does not inspire confidence in EEC, even where the correct identification is made, as the selection performed in such cases becomes an arbitrary choice between poor candidates.

The second problem arises where a number of similar and high valued correlation results are returned for a particular event edge in the total power measurements. In such cases, EEC is indicating that more than one appliance is closely matched to the captured edge sample and that it is difficult to determine which is the best fit. Where the difference between correlations may be as low as 0.001 in places, a correct identification may be considered to be a random outcome. However, as EEC still does limit the range of likely appliance matches in such cases, any randomness contained in the final assignment is reduced by the process of elimination that proceeds it. Nonetheless, confidence in EEC would be encouraged by an increase in the differences found between all of the correlation values returned for each event edge, allowing for more definitive appliance identifications to be made.

This concludes the presentation and discussion of the results and observations. Recommendations for further work in this research area are presented in Chapter 6.

Chapter 6

FUTURE WORK

Proposals for future work in the research area covered within this dissertion are presented and discussed. An alternative methodology for the measurements process is outlined. Further refinements of EEC that could improve the disaggregation accuracy of the technique are presented and discussed.

6.1 Future Work Overview

Three proposals for further work in this research area are presented in this chapter. The first of these applies to future measurements that may be performed, where a large number of appliance state combinations must be measured in conjunction with accurate ground truth data. The development of an automated measurement system for this purpose is proposed and discussed in Section 6.2.

The results and observations presented in Chapter 5 reveal EEC to be the

most suitable of the three disaggregation techniques for further development. Thus a refinement to EEC is proposed in Section 6.3, wherein it is suggested that the history of appliance identifications made earlier in a disaggregation pass could be used to improve the accuracy of the remaining event edge identifications via logical elimination.

As discussed in Section 5.6, any refinement of the sample extraction process that emphasises the distinctiveness of the waveform edges captured in the sample will result in improved disaggregation accuracy for EEC. Thus, whilst it is not discussed further in this chapter, the sample extraction process is also proposed as a potential avenue for future development of EEC. Advances made in this area could be applied to any similar transient disaggregation techniques that rely upon the distinctiveness of event edges in order to identify appliance state transitions.

6.2 Automated Measurement System

For the research presented in this dissertation, pre-planned combinations of appliance operations were enacted and measured in the laboratory, providing nine combined measurements that simulate possible real-world scenarios. The resulting total power measurements do not include every possible combination of appliances, only those specified in the ground truth data. It would be useful to be able to take measurements for every possible combination and sequence of appliance states, producing a far larger and richer database of ground truth data and corresponding measurements. This would allow for disaggregation techniques to be more thoroughly and methodically tested.

A system that might be developed in order to obtain such data is presented in Figure 6.1 and described in the following points.

1. An automated system must be devised that may be programmed to switch the operational states of the appliances included in the experiment in a se-

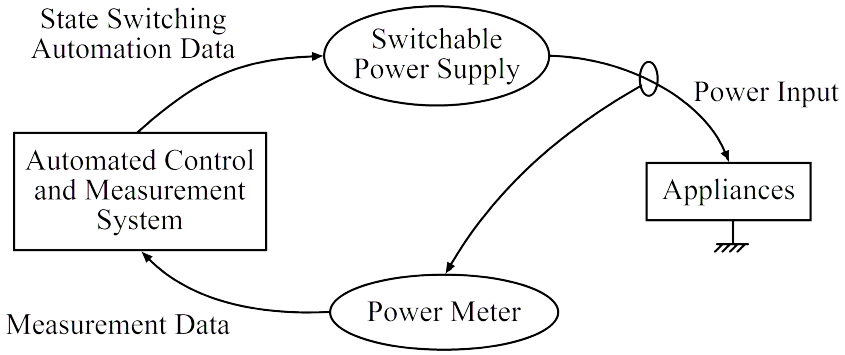

Figure 6.1: Automated appliance measurement system.

quence dictated by the researcher, whilst simultaneously producing accompanying ground truth data. Appropriate power measurements must be taken throughout this process, in a manner that allows them to be synchronised to the ground truth data at the end of each iteration of the experiment. Some of the appliances may have to be physically altered such that the automated system can directly control their operational states, or that manipulation of their power supply is sufficient to activate and deactivate them as required.

2. As with the research contained within this dissertation, individual measurements of all appliances included in the experiment must be taken by the system to supplement the combined measurements. Once captured, these individual measurements would have to be analysed by the researcher such that they could be compiled into an exemplar library, the requirements of which would depend upon the peculiarities of the disaggregation technique(s) to be investigated.

3. When combined measurements are taken, the system must run through every feasible combination of appliance operational states, activating each state for a fixed period of time before moving onwards to the next set of states. The sequence of events must be altered for each iteration of the experiment, such that the disaggregation technique to be applied to the gathered data may encounter different sequences of appliance states.

Producing a comparable data set by manual means would be a time intensive

undertaking, potentially leading to a low limit being placed upon the number of appliances included in the experiment and thus diminishing the value of any associated research. Given that the eight appliances utilised in the laboratory experiment presented in this dissertation produced 2048 possible appliance state combinations, the time required for even a single manual pass of the experiment covering all state combinations would be prohibitive. Furthermore, the length of the process would lead to the introduction of copious human errors, most likely appearing between the prescribed ground truth data and the actual sequences of appliance operational states being enacted. As researchers are unlikely to be able to perform appliance state transitions at sub-second accuracy rates, an automated system would also offer improved synchronisation between measurements and the corresponding ground truth data.

It would be advantageous to be able to test any disaggregation techniques of interest against several different iterations of the experiment, such that the consistency of performances may be evaluated. Implementation of this automated approach would allow for accurate, tightly synchronised, extensive and repeatable data sets to be relatively easily generated. Whilst the measurement process followed in this research does simulate the real-world in many ways, it repeats and excludes certain states and sequences, as discussed in Chapter 5. Although the inconsistencies observed over repeated states within the existing measurements are of interest in themselves, it would be preferable to be able to sample every feasible combination of devices and state change sequences, given that repeated states could certainly be included in the ground truth data plan where desired.

6.3 Previous Identifications

EEC stores no memory of previous appliance state identifications, as discussed in Section 5.4. If such functionality were to be included, it would introduce an element of logic into the technique, whereby infeasible state

changes could be excluded from the comparison process. This would serve to improve accuracy by reducing the number of appliance edge samples available to be matched to each event edge in the total power measurement. For example, if an appliance positive edge has not been identified during the course of a measurement, then its negative edge may not be considered for comparison to any negative edges detected in the combined measurement. This approach assumes that all appliances are inactive prior to the start of the process. Or if an appliance has previously been determined to have become active, and has not yet been deactivated, then its positive edge could not be considered for matching to any positive edges detected in the total power measurement. Under the correct conditions, such a scheme of logical elimination could improve disaggregation accuracy.

However, the proposed system would rely heavily on the accuracy of previous identifications in order to proceed without significant error. Where a disaggregation error is made, all subsequent identifications will be affected by the erroneous application of the resulting logical elimination, until such point that some fortunate series of appliance state changes allows for the original error to be discarded. For example, if an appliance is erroneously identified as becoming operational, all of the following event edge comparisons will exclude that appliance even though it may be responsible. Thus, the original error would be perpetuated through the disaggregation pass until the incorrectly identified appliance becomes active, or the memory of previous identifications is cleared. The use of previous identifications to supplement EEC is of considerable interest, most definitely offering the potential to increase its accuracy. However, it would be best implemented in the context of improved performance of the principle mechanics of the technique, where EEC has already been refined further using other methods.

This concludes the presentation and discussion of future work. The conclusion of the dissertation, which draws upon all of the work contained within the preceding chapters of this dissertation, is presented in Chapter 7.

Chapter 7

CONCLUSION

Three disaggregation techniques are comparatively evaluated in this dissertation, namely; Total Load Model (TLM), Complex Load Method (CPM) and Event Edge Correlation (EEC). ECC offers the best performance, with an overall hard score accuracy of 52%, and soft score accuracy of 57%. CPM achieved the next best performance, with an overall accuracy of 40%, and TLM exhibited the worst performance of the three disaggregation techniques, with an overall accuracy of 24%. All three of these accuracies were found using the equal weighting approach, with the outlier combined measurements 5 and 7 excluded, as discussed in Section 5.5. If these two combined measurements are included in the results, then the difference between EEC and CPM is more marginal; EEC offers a hard score accuracy of 47% and a soft score accuracy of 57%, CPM exhibits an accuracy of 46%, and TLM again offers the worst accuracy at 31%. Motivation for the exclusion of the outlier combined measurements 5 and 7 is given in Section 5.5.

These results may be generalised to say that the comparison of transient event edges in the total power measurement provides a better disaggregation approach than that of steady-state power levels. Furthermore, the consideration of both real and reactive power data in the context of steady-state power disaggregation may be considered to be superior to the exclusive use of real

power data. Whilst EEC does not achieve sufficient levels of accuracy to be directly implemented in a real-world NILM system, it should be considered ahead of TLM and CPM for further development, which could refine it to the point where it may deliver excellent accuracies. This is especially the case for circuits containing larger numbers of appliances, as EEC stands to suffer the least degradation of performance as the number of appliances included in the measured circuit is increased.

Five key observations arise from analysis of the results. The first observation is that the high number of similar valued power consumption levels found for different appliance state combinations has an adverse effect on the performance of TLM and CPM, which is consistent with the expected challenges listed in Section 1.4. The second observation is that the degree of distinctiveness of appliance waveform event edges has a direct positive effect on the ability of EEC to distinguish them from those of other appliances, and hence on the accuracy of the technique. This is an extension of the same expected challenge mentioned for the first observation.

The third observation is the existence of favourable regions within the appliance state combination vector, where the probability of making accurate matches is higher than for other regions of the vector. This makes it easier to obtain accurate appliance identifications for some combinations of appliance states than for others, depending on the position of the measured power consumption level relative to the favourable matching regions. This is consistent with the expected challenges presented by combinatorial approaches, as discussed in Section 1.4.

The fourth observation is that the variability found in measured power levels affects the accuracy of all three of the disaggregation techniques. A combination of variable appliance operations, hidden appliance states and the presence of electrical noise, cause discrepancies between measured and expected power levels, introducing error into the appliance state identifications. This observation is consistent with the expected challenges discussed in Section 1.4. Whilst TLM and CPM are more adversely affected by this factor than

EEC, variations in measured power levels can alter the shapes of appliance waveform edges, and thus the accuracy of transient disaggregation techniques such as EEC.

The final observation is that many of the correlation results returned by EEC for each event edge in the combined measurements do not present definitive identifications. In many cases, only low correlation values are produced for an event edge, and thus the selected edge is not a good match to the event edge in the power measurement, but rather the best out of a collection of poor matches. Alternatively, a number of the correlation values produced for a single event edge may indicate almost perfect correlation, and thus any selection made between such values is not definitive. This observation did not form part of the expected challenges, making it a new area for consideration for EEC.

The results and observations lead to recommendations for future work to be conducted in the same research area. An alternative measurements methodology is proposed, involving the development of an automated measurement system that would be capable of running through large combinations of appliance operational states with high accuracy, whilst compiling associated ground truth data. Such an automated system would be able to perform measurements over extended time periods, without human error compromising the accuracy of the ground truth reporting, producing a definitive dataset for the evaluation of total power disaggregation techniques.

Two further recommendations may be made, both of which relate to improving the accuracy of EEC and other similar transient disaggregation techniques. The first is that the sample extraction process be refined, such that the characteristic shape of the features contained within the samples produced for comparison are emphasised. This will increase their distinctiveness, and hence the accuracy with which EEC identifies them. The second involves the use of previous identifications to refine the matching process via logical elimination. However, the underlying accuracy of EEC would need to be improved before the introducing such a system.

REFERENCES

[1] P. Palensky and D. Dietrich. "Demand Side Management: Demand Response, Intelligent Energy Systems, and Smart Loads." *IEEE Transactions on Industrial Informatics*, vol. 7, no. 3, pp. 381–388, Aug. 2011.

[2] G. W. Hart. "Nonintrusive Appliance Load Monitoring." *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec. 1992.

[3] "MathWorks - MATLAB and Simulink for Technical Computing." `http://http://www.mathworks.com`. Accessed: 2015-08-21.

[4] M. Berges, E. Goldman, S. Matthews, L. Soibelman, and K. Anderson. "User-Centred Nonintrusive Electricity Load Monitoring for Residential Buildings." *Journal of Computing in Civil Engineering*, vol. 25, no. 6, pp. 471–480, Nov. 2011.

[5] M. Figueiredo, A. de Almeida, and B. Ribeiro. "Home Electrical Signal Disaggregation for Non-Intrusive Load Monitoring (NILM) Systems." *Neurocomputing*, vol. 96, no. 1, pp. 66–73, Nov. 2012.

[6] S. B. Leeb, S. R. Shaw, and J. L. Kirtley. "Transient Event Detection in Spectral Envelope Estimates for Nonintrusive Load Monitoring." *IEEE Transactions on Power Delivery*, vol. 10, no. 3, pp. 1200–1210, Jul. 1995.

[7] S. R. Shaw, S. B. Leeb, L. K. Norford, and R. W. Cox. "Nonintrusive Load Monitoring and Diagnostics in Power Systems." *IEEE Transactions on Instrumentation and Measurement*, vol. 57, no. 7, pp. 1445–1454, Jul. 2008.

[8] G. W. Hart. "Residential Energy Monitoring and Computerized Surveillance Via Utility Power Flows." *IEEE Technology and Society Magazine*, vol. 8, no. 2, pp. 12–16, Jun. 1989.

[9] M. Dong, P. C. M. Meira, W. Xu, and W. Freitas. "An Event Window Based Load Monitoring Technique for Smart Meters." *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 787–796, Jun. 2012.

[10] F. Sultanem. "Using Appliance Signatures for Monitoring Residential Loads at Meter Panel Level." *IEEE Transactions on Power Delivery*, vol. 6, no. 4, pp. 1380–1385, Oct. 1991.

[11] M. L. Marceau and R. Zmeureanu. "Nonintrusive Load Disaggregation Computer Program to Estimate the Energy Consumption of Major End Uses in Residential Buildings." *Energy Conversion and Management*, vol. 41, no. 13, pp. 1389–1403, Sep. 2000.

[12] A. I. Cole and A. Albicki. "Data Extraction for Effective Non-Intrusive Identification of Residential Power Loads." In *Proceedings of the 1998 IEEE Instrumentation and Measurement Technology Conference*, pp. 812–815. St. Paul, May 1998.

[13] H. Najmeddine, K. E. E. K. Drissi, C. Pasquier, C. Faure, K. Kerroum, A. Diop, T. Jouannet, and M. Michou. "State of Art on Load Monitoring Methods." In *2nd IEEE International Conference on Power and Energy*, pp. 1256–1258. Johor Baharu, Dec. 2008.

[14] L. K. Norford and S. B. Leeb. "Non-Intrusive Electrical Load Monitoring in Commercial Buildings Based on Steady-State and Transient Load-Detection Algorithms." *Energy and Buildings*, vol. 24, no. 1, pp. 51–64, Jan. 1996.

[15] Y. Jin, E. Tebekaemi, M. Berges, and L. Soibelman. "Robust Adaptive Event Detection in Non-Intrusive Load Monitoring for Energy Aware Smart Facilities." In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 4340–4343. Prague, May 2011.

[16] K. D. Anderson, M. E. Berges, A. Ocneanu, D. Benitez, and J. M. F. Moura. "Event Detection for Non-Intrusive Load Monitoring." In *38th Annual Conference on IEEE Industrial Electronics Society*, pp. 3312–3317. Montreal, Oct. 2012.

[17] C. Laughman, K. Lee, R. Cox, S. Shaw, S. Leeb, L. Norford, and P. Armstrong. "Power Signature Analysis." *IEEE Power and Energy Magazine*, vol. 1, no. 2, pp. 56–63, Mar. 2003.

[18] C. Duarte, P. Delmar, K. W. Goossen, K. Barner, and E. Gomez-Luna. "Non-Intrusive load Monitroing Based on Switching Voltage Transients and Wavelet Transforms." In *2012 Future of Instrumentation International Workshop*, pp. 1–4. Gatlinburg, Oct. 2012.

[19] A. I. Cole and A. Albicki. "Algorithm for Non-Intrusive Identification of Residential Appliances." In *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, pp. 338–341. Monterey, Jun. 1998.

[20] H. H. Chang and C. L. Lin. "A New Method for Load Identification of Nonintrusive Energy Management System in Smart Home." In *2010 IEEE 7th International Conference on e-Business Engineering*, pp. 351–357. Shanghai, Nov. 2010.

[21] T. Saitoh, Y. Aota, T. Osaki, R. Konishi, and K. Sugahra. "Current Sensor Based Non-Intrusive Appliance Recognition for Intelligent Outlet." In *Proceedings of the 23rd International Technical Conference on Circuits/Systems, Computers and Communications*, pp. 349–352. Shimonoseki, Jul. 2008.

[22] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng. "Load Signature Study - Part I: Basic Concept, Structure, and Methodology." *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, Apr. 2010.

[23] H. S. Matthews, L. Soibelman, M. Berges, and E. Goldman. "Automatically Disaggregating the Total Electrical Load in Residential Buildings: a Profile of the Required Solution." In *15th International Workshop*

on Intelligent Computing in Engineering, pp. 381–389. Plymouth, Jul. 2008.

[24] Y. Du, L. Du, R. Harley, and T. Habetler. "A Review of Identification and Monitoring Methods for Electric Loads in Commercial and Residential Buildings." In *2010 IEEE Energy Conversion Congress and Exposition*, pp. 4527–4533. Atlanta, Sep. 2010.

[25] L. Farinaccio and R. Zmeureanu. "Using a Pattern Recognition Approach to Disaggregate the Total Electricity Consumption in a House into the Major End-Uses." *Energy and Buildings*, vol. 30, no. 3, pp. 245–259, Aug. 1999.

[26] H. Y. Lam, G. S. K. Fung, and W. K. Lee. "A Novel Method to Construct Taxonomy of Electrical Appliances Based on Load Signatures." *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 653–660, May 2007.

[27] M. Zeifman and K. Roth. "Nonintrusive Appliance Load Monitoring." *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 76–84, Feb. 2011.

[28] K. Anderson, A. Ocneanu, D. Benitez, D. Carlson, A. Rowe, and M. Berges. "BLUED: A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research." In *Proceedings of the 2nd KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*, pp. 1–5. Beijing, Aug. 2012.

[29] J. Z. Kolter and M. J. Johnson. "REDD: A Public Data Set for Energy Disaggregation Research." In *Proceedings of the 1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD)*. San Diego, Aug. 2011.

[30] S. Makonin, F. Popowich, L. Bartram, B. Gill, and I. V. Bajic. "AMPds: A Public Dataset for Load Disaggregation and Eco-Feedback Research." In *Proceedings of the 2013 IEEE Electrical Power and Energy Conference (EPEC)*, pp. 1–6. Halifax, Aug. 2013.

[31] Y. M. Corporation. *CW240 Clamp-on Power Meter User's Manual*. Tokyo, first ed., 2004.

[32] Eskom. "Standard Conditions of Supply for Large Supplies." *www.eskom.co.za*, pp. 1–10, Nov. 2007.

[33] C. G. Carter-Brown and C. T. Gaunt. "Distribution Voltage Regulation Management and Optimisation." In *24th Association of Municipal Electricity Utilities of Southern Africa Technical Convention 2013*, pp. 1–9. Buffalo City, Oct. 2013.

[34] M. Weiss, A. Helfensteni, F. Mattern, and T. Staake. "Leveraging Smart Meter Data to Recognise Home Appliances." In *2012 IEEE International Conference on Pervasive Computing and Communications*, pp. 190–197. Lugano, Mar. 2012.

[35] "MathWorks - Documentation: 'Corr' Linear or Rank Correlation." `http://http://www.mathworks.com`. Accessed: 2015-06-14.

[36] N. S. Chok. *Pearson's versus Spearman's and Kendall's Correlation Coefficients for Continuous Data*. Master's thesis, Winona State University, Winona, 2010.

Appendix A

REACTIVE POWER MEASUREMENT PLOTS

This appendix contains reactive power plots that correspond to discussions conducted in Chapter 3. Both individual and combined measurement reactive power plots are included, in order to supplement the real power plots contained within Chapter 3. Where appliances possess a reactive power component, it may be utilised as an additional layer of information to aid in the identification of individual appliance operations within total power measurements. This is the case with CPM, where steady-state reactive power levels are used to refine the disaggregation process.

As may be seen in the body of the appendix, the majority of the appliances utilised in the laboratory experiment do not possess a reactive power component. This is reflected in the combined measurement plots, where any reactive power waveforms may be ascribed to the operation of those appliances that do posses a reactive power component, namely; the fridge and microwave. Throughout the plots, capacitive and inductive transient peak events may be observed. These transients accompany appliance operational state switching events, as discussed within Chapter 3.

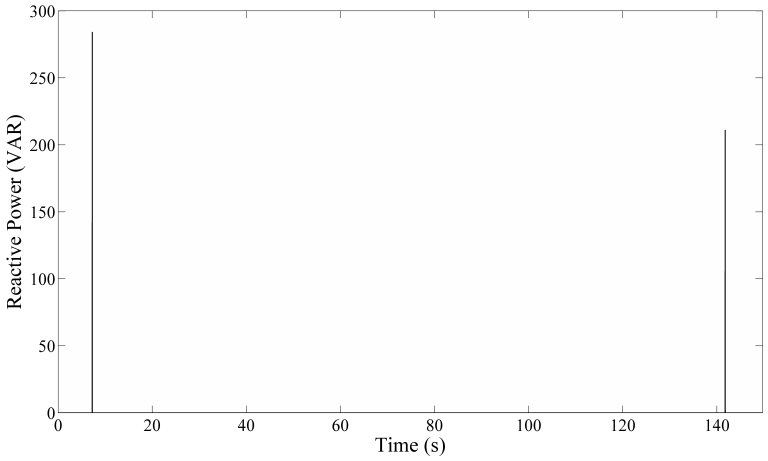

Figure A.1: Reactive power waveform for toaster.

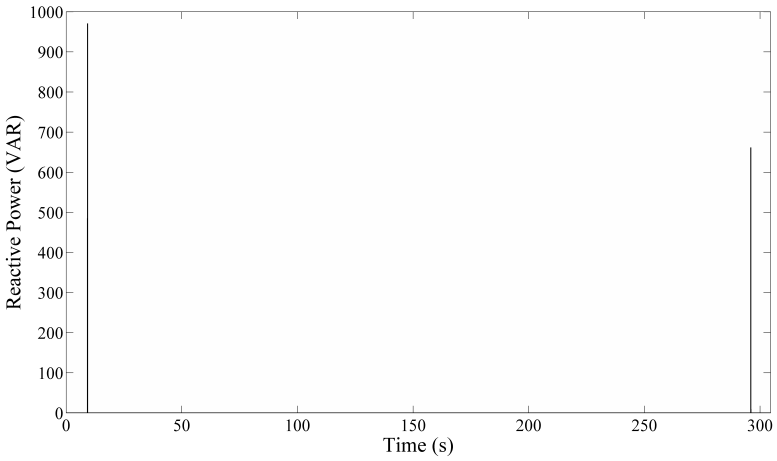

Figure A.2: Reactive power waveform for kettle.

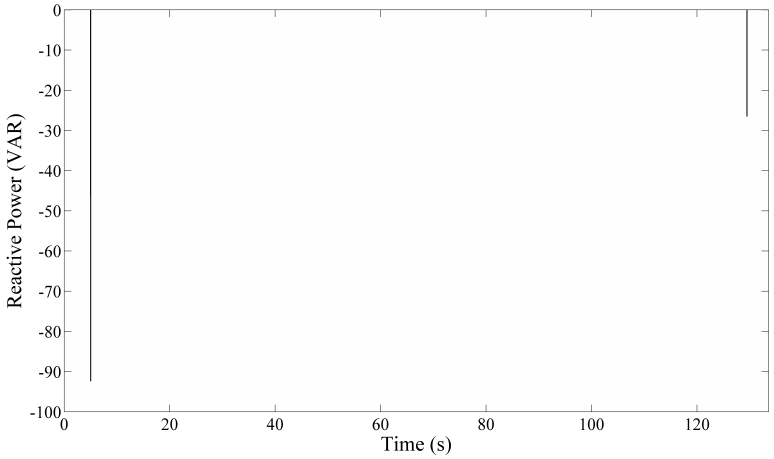

Figure A.3: Reactive power waveform for lamp.

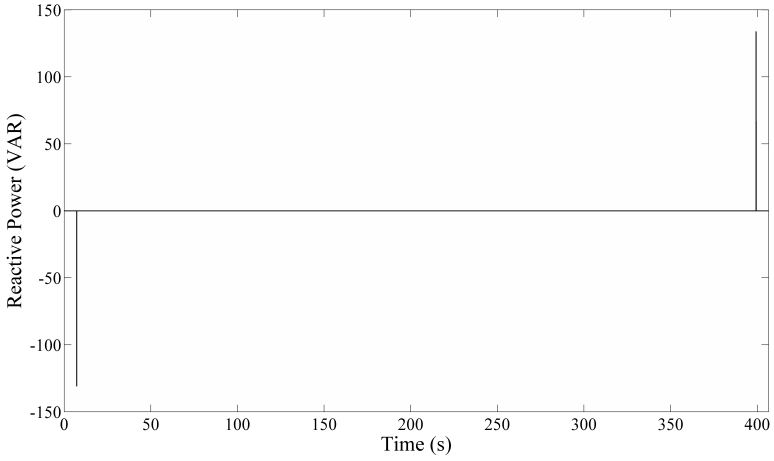

Figure A.4: Reactive power waveform for heater-low.

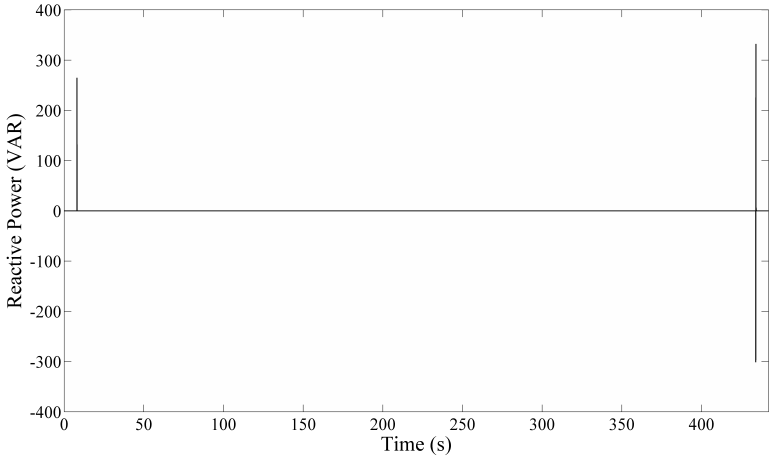

Figure A.5: Reactive power waveform for heater-medium.

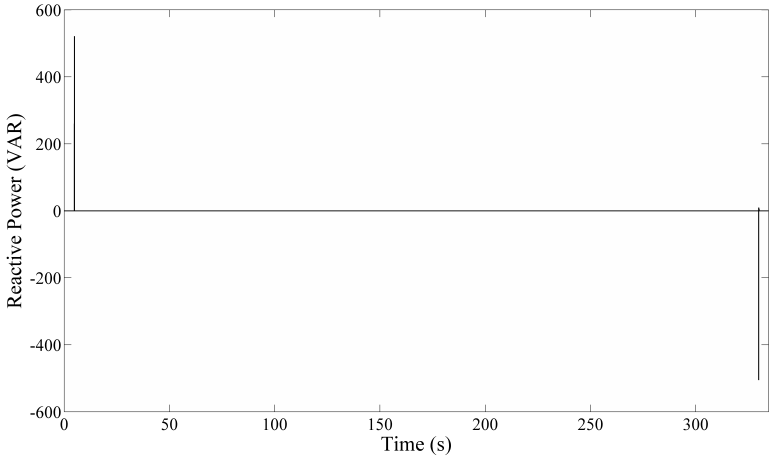

Figure A.6: Reactive power waveform for heater-high.

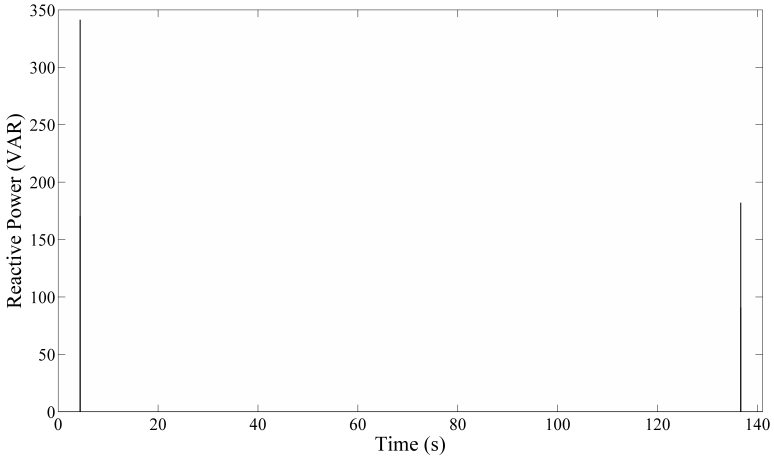

Figure A.7: Reactive power waveform for snackwich.

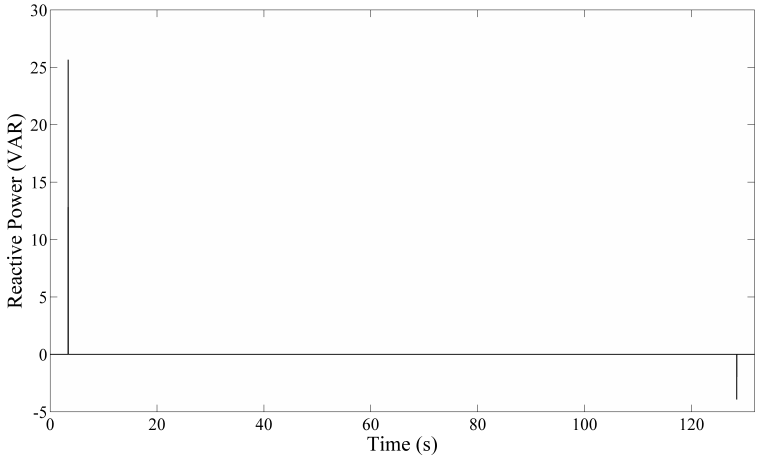

Figure A.8: Reactive power waveform for fan-low.

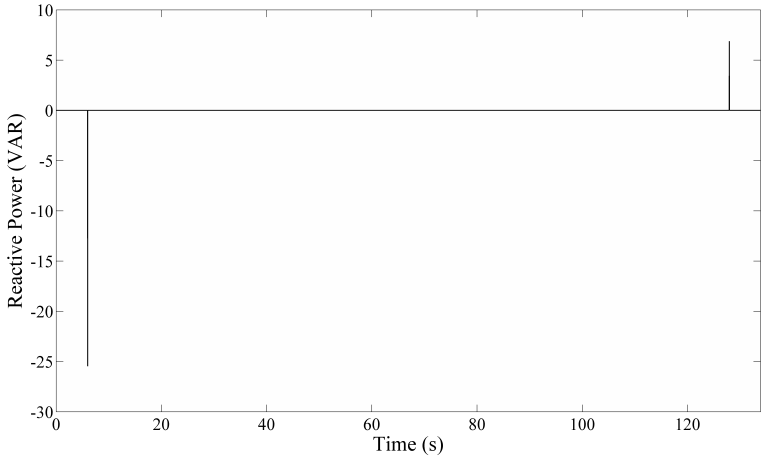

Figure A.9: Reactive power waveform for fan-medium.

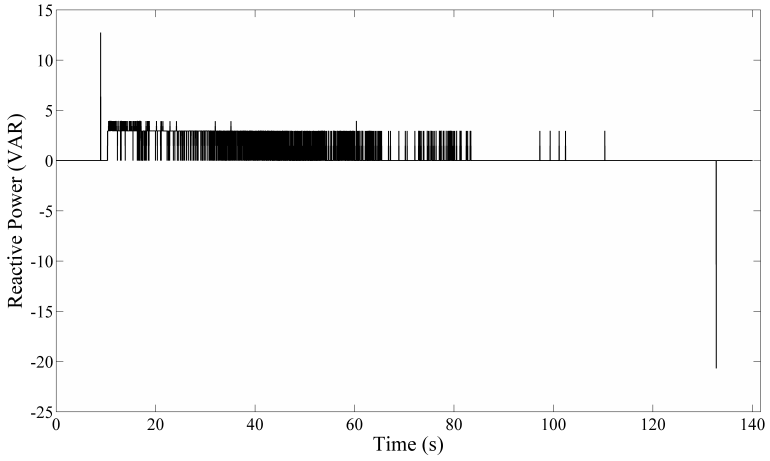

Figure A.10: Reactive power waveform for fan-high.

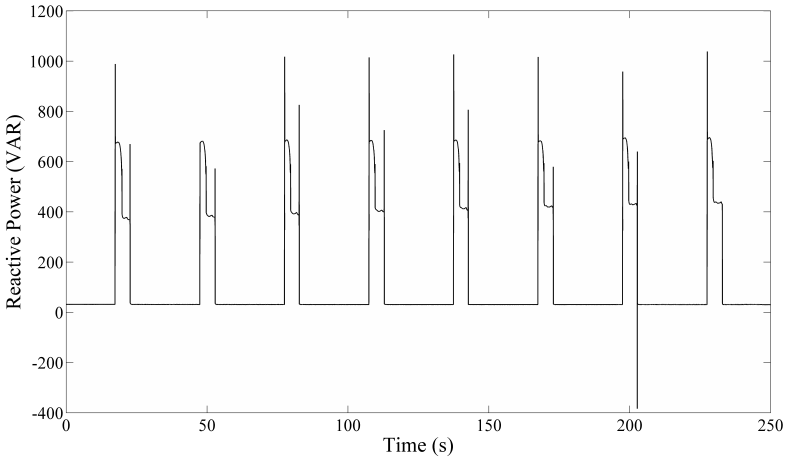

Figure A.11: Reactive power waveform for microwave low power cooking setting.

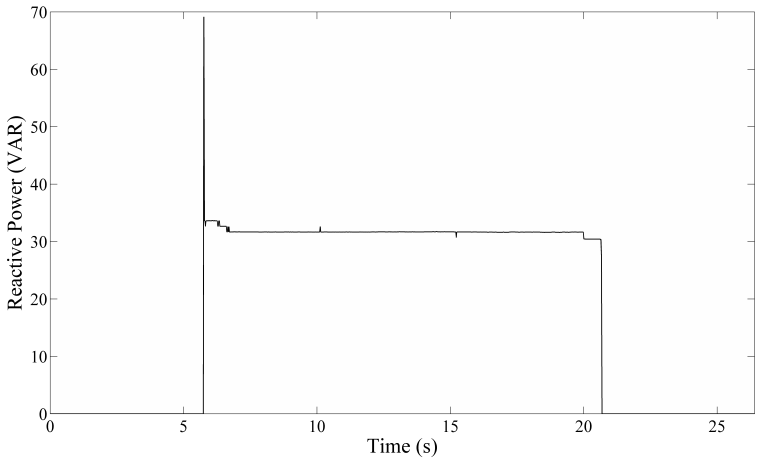

Figure A.12: Reactive power waveform for microwave-inactive.

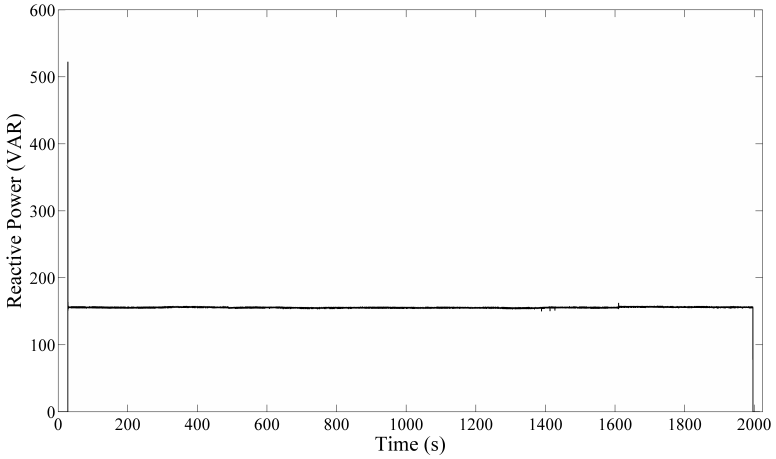

Figure A.13: Reactive power waveform for fridge.

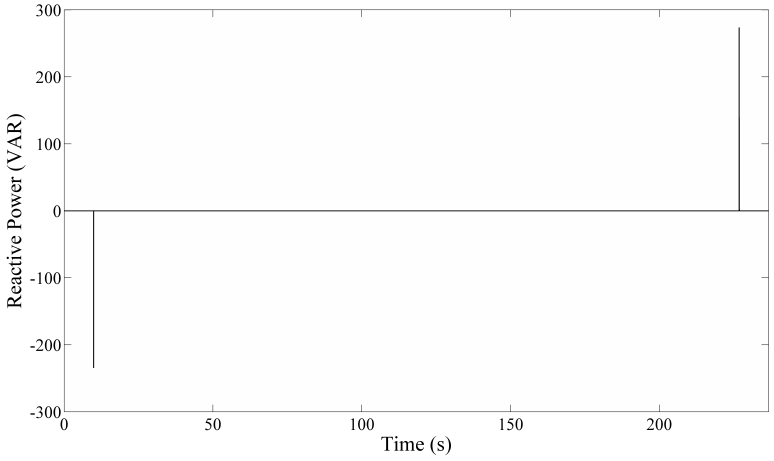

Figure A.14: Combined Measurement 1 reactive power plot.

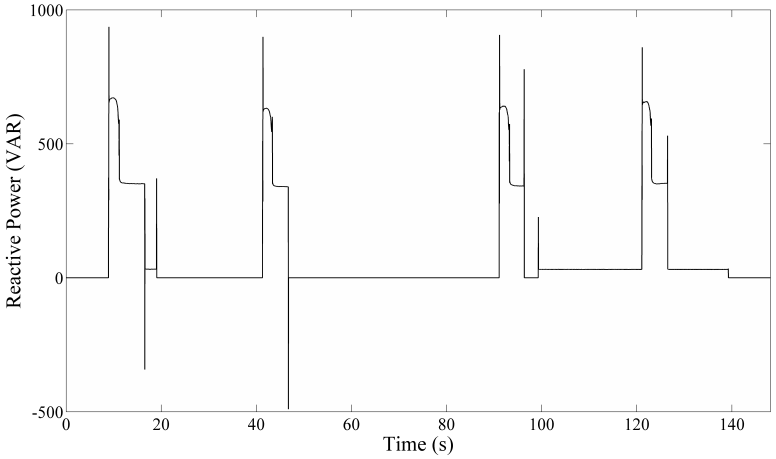

Figure A.15: Combined Measurement 2 reactive power plot.

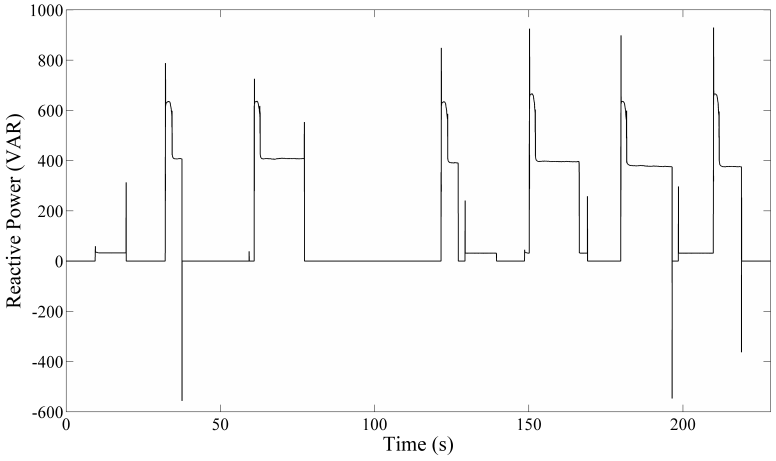

Figure A.16: Combined Measurement 3 reactive power plot.

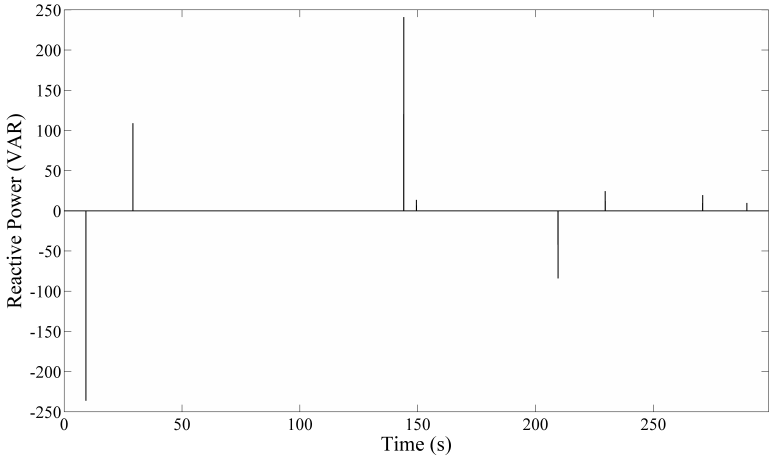

Figure A.17: Combined Measurement 4 reactive power plot.

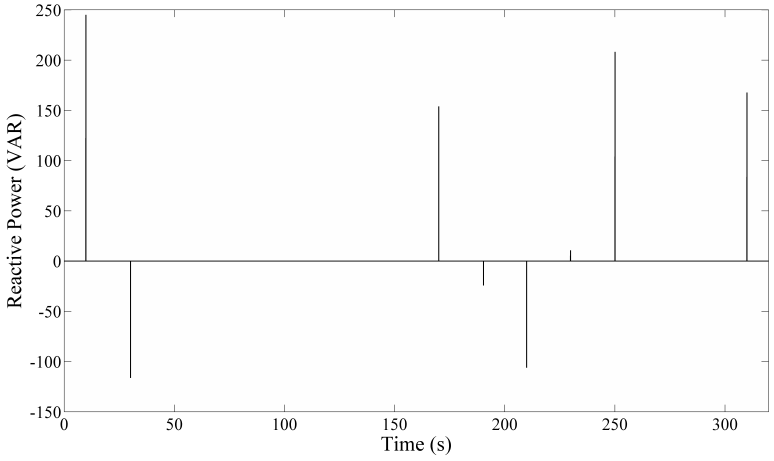

Figure A.18: Combined Measurement 5 reactive power plot.

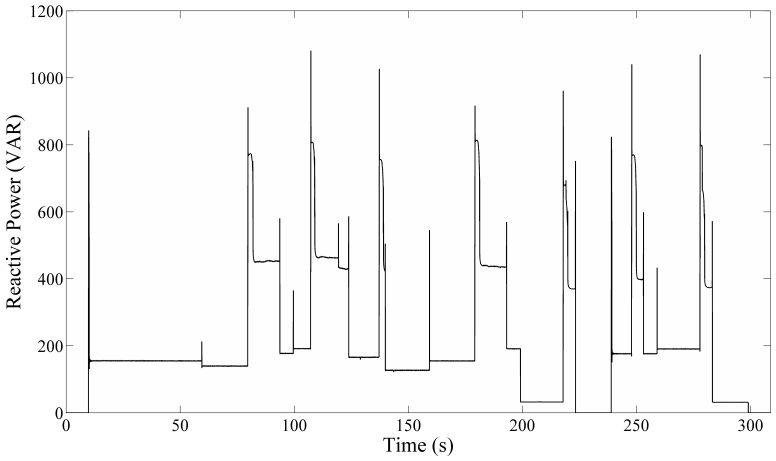

Figure A.19: Combined Measurement 6 reactive power plot.

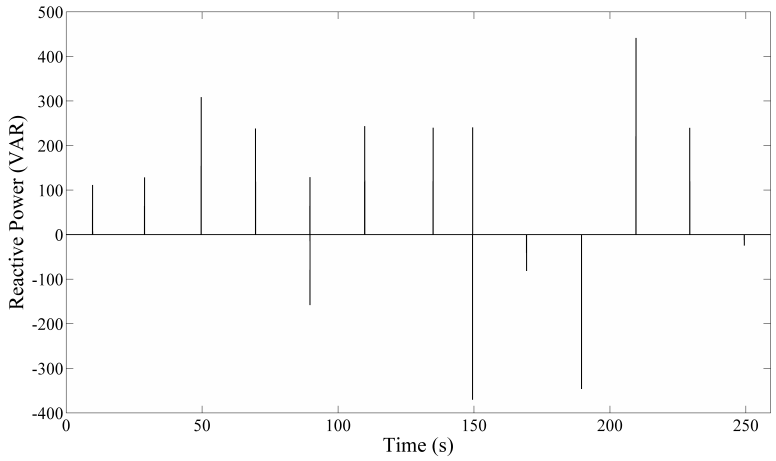

Figure A.20: Combined Measurement 7 reactive power plot.

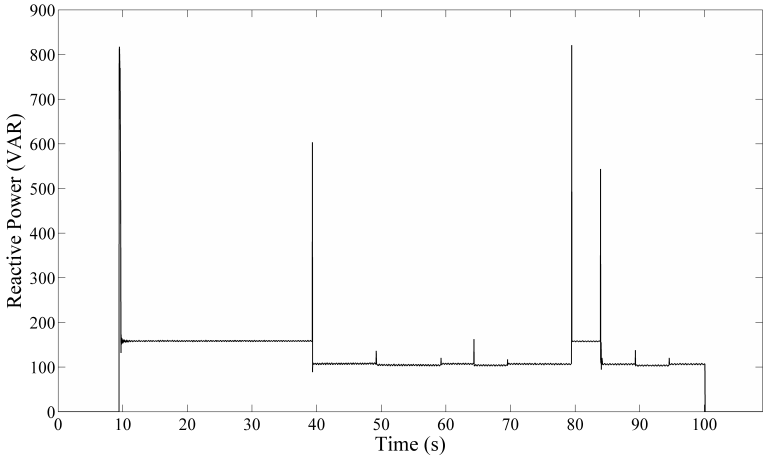

Figure A.21: Combined Measurement 8 reactive power plot.

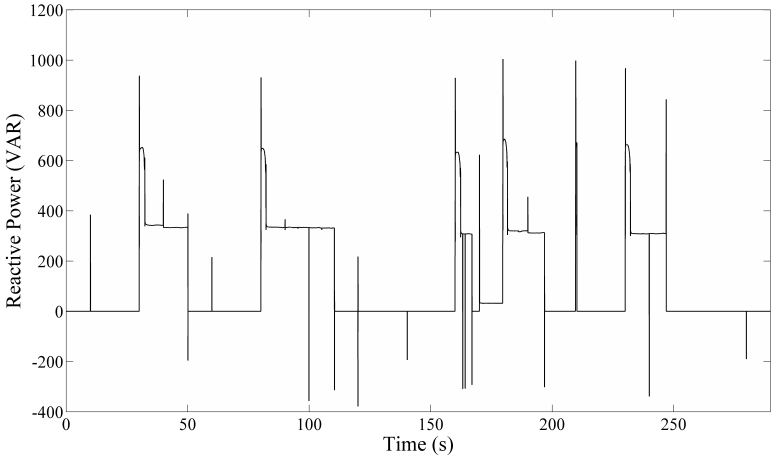

Figure A.22: Combined Measurement 9 reactive power plot.

Appendix B

GROUND TRUTH DATA

This appendix contains tables of ground truth data for the combined measurements, where Tables B.1 through B.9 correspond to each of the nine combined measurements respectively. As some appliance actions are automated, not all entries originally possessed ground truth data time stamps. These additions to the underlying time-stamped ground truth data were made via inspection of the measurements during post-processing.

The use of accurate ground truth data, where researchers are precisely aware of the exact sequences of appliance operational state change events, allows for the effective evaluation of total power measurement disaggregation techniques. Thus, the data contained within this appendix is fundamental to this research, as it presents the true sequence of appliance state change events to which TLM, CPM and EEC were applied.

A legend to the appliance name abbreviations utilised in the tables contained within this appendix is provided below.

Appliance	**Abbreviation**
Toaster	tst
Kettle	ktl
Lamp	lmp
Heater-Low	ht1
Heater-Medium	ht2
Heater-High	ht3
Snackwich	snw
Fan-Low	fnL
Fan-Medium	fnM
Fan-High	fnH
Microwave-Front	mwF
Microwave-Back	mwB
Microwave-Inactive	mwI
Microwave-Low	mwL
Microwave-Medium	mwM
Microwave-High	mwH
Fridge	frg

Table B.1: Combined Measurement 1 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	tst ON	tst
20	fnL ON	tst,fnL
30	fnL OFF	tst
40	fnM ON	tst,fnM
50	fnM OFF	tst
60	fnH ON	tst,fnH
70	fnH OFF	tst
227	tst OFF	null

Table B.2: Combined Measurement 2 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	mwL ON, mwF AUTO	mwF
–	mwB AUTO	mwB
–	mwI AUTO	mwI
20	tst ON	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
60	mwL OFF	tst
80	mwL ON, mwI AUTO	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
100	tst OFF	mwI
–	mwF AUTO	mwF
–	mwB AUTO	mwB
–	mwI AUTO	mwI
140	mwL OFF	null

Table B.3: Combined Measurement 3 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	mwL ON, mwI AUTO	mwI
20	tst ON	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
50	mwL OFF	tst
60	mwM ON, mwI AUTO	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
90	mwM OFF	tst
110	mwL ON, mwI AUTO	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
130	tst OFF	mwI
140	mwL OFF	null
150	mwM ON, mwI AUTO	mwI
–	mwF AUTO	mwF
–	mwB AUTO	mwB
–	mwI AUTO	mwI
170	tst ON	tst,mwI
–	mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
–	mwI AUTO	tst,mwI
200	tst OFF	mwI
–	mwF AUTO	mwF
–	mwB AUTO	mwB
220	mwM OFF	null

Table B.4: Combined Measurement 4 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	snw ON	snw
30	lmp ON	snw,lmp
50	fnL ON	snw,lmp,fnL
70	fnL OFF	snw,lmp
90	fnM ON	snw,lmp,fnM
110	fnM OFF	snw,lmp
130	fnH ON	snw,lmp,fnH
144	snw OFF	lmp,fnH
150	fnH OFF	lmp
190	lmp OFF	null
210	lmp ON	lmp
230	fnM ON	lmp,fnM
270	lmp OFF	fnM
290	fnM OFF	null

Table B.5: Combined Measurement 5 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	ht1 ON	ht1
30	lmp ON	ht1,lmp
50	fnL ON	ht1,lmp,fnL
70	fnL OFF	ht1,lmp
90	fnM ON	ht1,lmp,fnM
110	fnM OFF	ht1,lmp
130	fnH ON	ht1,lmp,fnH
150	fnH OFF	ht1,lmp
170	ht1 OFF	lmp
190	lmp OFF	null
210	lmp ON	lmp
230	fnM ON	lmp,fnM
250	ht1 ON	lmp,fnM,ht1
270	lmp OFF	fnM,ht1
290	fnM OFF	ht1
310	ht1 OFF	null

Table B.6: Combined Measurement 6 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	frg ON	frg
60	ht2 ON	frg,ht2
80	mwM ON, mwF AUTO	frg,ht2,mwF
–	mwB AUTO	frg,ht2,mwB
–	mwI AUTO	frg,ht2,mwI
100	ht2 OFF	frg,mwI
–	mwF AUTO	frg,mwF
–	mwB AUTO	frg,mwB
120	ht3 ON	ht3,frg,mwB
–	mwI AUTO	frg,ht3,mwI
–	mwF AUTO	frg,ht3,mwF
–	mwB AUTO	frg,ht3,mwB
140	mwM OFF	frg,ht3
160	ht3 OFF	frg
180	mwL ON, mwF AUTO	frg,mwF
–	mwB AUTO	frg,mwB
–	mwI AUTO	frg,mwI
200	frg OFF	mwI
–	mwF AUTO	mwF
220	ht2 ON, mwB AUTO	ht2,mwB
–	mwI AUTO	ht2,mwI
240	frg ON	frg,ht2,mwI
–	mwF AUTO	frg,ht2,mwF
–	mwB AUTO	frg,ht2,mwB
–	mwI AUTO	frg,ht2,mwI
260	ht2 OFF	frg,mwI
–	mwF AUTO	frg,mwF
280	frg OFF	mwF
–	mwB AUTO	mwB
–	mwI AUTO	mwI
300	mwL OFF	null

Table B.7: Combined Measurement 7 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	lmp ON	lmp
30	ht1 ON	lmp,ht1
50	tst ON	lmp,ht1,tst
70	ht1 OFF	lmp,tst
90	ht2 ON	lmp,ht2,tst
110	ht2 OFF	lmp,tst
130	lmp OFF	tst
135	tst OFF	null
150	ht3 ON	ht3
170	lmp ON	ht3,lmp
190	tst ON	ht3,lmp,tst
210	ht3 OFF	lmp,tst
230	tst OFF	lmp
250	lmp OFF	null

Table B.8: Combined Measurement 8 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	frg ON	frg
40	ktl ON	frg,ktl
50	lmp ON	frg,ktl,lmp
60	lmp OFF	frg,ktl
65	lmp ON	frg,ktl,lmp
70	lmp OFF	frg,ktl
80	ktl OFF	frg
85	ktl ON	frg,ktl
90	lmp ON	frg,ktl,lmp
95	lmp OFF	frg,ktl
100	frg OFF	ktl

Table B.9: Combined Measurement 9 ground truth data.

Time (s)	**State Switching Event**	**SS Combination Following Event**
10	tst ON	tst
30	mwH ON, mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
40	ht2 ON	tst,mwB,ht2
50	mwH OFF	tst,ht2
60	ht2 OFF	tst
80	mwH ON, mwF AUTO	tst,mwF
–	mwB AUTO	tst,mwB
90	lmp ON	tst,mwB,lmp
95	lmp OFF	tst,mwB
100	lmp ON	tst,mwB,lmp
105	lmp OFF	tst,mwB
110	mwH OFF	tst
120	ht3 ON	tst,ht3
140	tst OFF	ht3
160	mwM ON, mwF AUTO	ht3,mwF
–	mwB AUTO	ht3,mwB
–	mwI AUTO	ht3,mwI
170	ht3 OFF	mwI
–	mwF AUTO	mwF
–	mwB AUTO	mwB
190	ht1 ON	mwB,ht1
–	mwI AUTO	mwI,ht1
210	mwM OFF	ht1
230	mwL ON, mwF AUTO	ht1,mwF
–	mwB AUTO	ht1,mwB
240	lmp ON	ht1,mwB,lmp
–	mwI AUTO	ht1,mwI,lmp
260	lmp OFF	ht1,mwI
270	mwL OFF	ht1
280	ht1 OFF	null

Appendix C

MATLAB CODE

This appendix contains a selection of MATLAB functions for the implementation of TLM, CPM and EEC. Only code excerpts deemed to be relevant to the implementation discussion have been included, as referenced from the text in Chapter 4. Comments included in the code provide additional information pertaining to the functioning of each excerpt, and to the overall MATLAB implementation system developed.

TLM, CPM and EEC were implemented entirely in a post-processing context. Once the measurements process had been completed, the resulting data was captured into data tables, upon which each of the disaggregation techniques could be applied in the form of MATLAB functions.

Appendices C.1 and C.2 contain general measurements processing functions that are used by multiple disaggregation techniques, as discussed in Section 4.2. Appendices C.3 to C.5 are specific to TLM, Appendices C.6 and C.7 to CPM, and Appendices C.8 to C.12 to EEC.

```
%declare variables
edgeDetP = zeros(length(P),1);
edgeDetPPos = zeros(length(P),1);
edgeDetPNeg = zeros(length(P),1);
edgeDetMask = 50; %minimum event gap mask
posEdgeVal = 15; negEdgeVal = 15; %sets bottom thresholds
arbLevel = 50; %arbitrary non-zero value used to denote edges

%replace 'NaN' values in real power signal with zeroes
noNaN = find(isnan(P)); P(noNaN) = 0;

for n = 1:edgeDetMask %applies minimum event gap mask
   edgeDetP(n) = NaN; edgeDetPPos(n) = NaN; edgeDetPNeg(n) ...
      = NaN;
end

for n = (edgeDetMask+1):length(P) %detects and assigns edges
    diff = P(n) - P(n-1);
    if (all(isnan(edgeDetP(n-1:-1:n-edgeDetMask))))
        if (diff > posEdgeVal)
            edgeDetP(n) = arbLevel; edgeDetPPos(n) = ...
                arbLevel; edgeDetPNeg(n) = NaN;
        elseif (diff < -negEdgeVal)
            edgeDetP(n) = arbLevel; edgeDetPNeg(n) = ...
                arbLevel; edgeDetPPos(n) = NaN;
        else
            edgeDetP(n) = NaN; edgeDetPPos(n) = NaN; ...
                edgeDetPNeg(n) = NaN;
        end
    else
        edgeDetP(n) = NaN; edgeDetPPos(n) = NaN; ...
            edgeDetPNeg(n) = NaN;
    end
end
```

Figure C.1: `eventDetection` function.

```
%call event detection function
eventDetection;

%declare variables and append '1' to front of vector
indSSP = find(~isnan(edgeDetP));
indSSP(2:end+1) = indSSP; indSSP(1) = 1;
ssDetP = zeros(length(P),1);
ssDetQ = zeros(length(P),1);
ssDetResP = zeros(length(indSSP),2);
ssDetResQ = zeros(length(indSSP),2);

%average values to find steady-state
for i = 1:(length(indSSP)-1)
    ssDetP(indSSP(i):indSSP(i+1)) = ...
        mean(P(indSSP(i):indSSP(i+1)));
    ssDetQ(indSSP(i):indSSP(i+1)) = ...
        mean(Q(indSSP(i):indSSP(i+1)));
end

%populate real and reactive power steady-state vectors
for i = 1:length(indSSP)
    ssDetResP(i,1) = indSSP(i);
    ssDetResP(i,2) = ssDetP(indSSP(i)+3);

    ssDetResQ(i,1) = indSSP(i);
    ssDetResQ(i,2) = ssDetQ(indSSP(i)+3);
end
```

Figure C.2: `ssTransformation` function.

```
%call alphabetical combinator function
TLMCharGenerator;

%declare variables
loopSize = 0;
endSize = 0;
catSize = 0;

%declare appliance state combination vectors
initVec{1} = [initMx(1) initMx(2) initMx(3)];
initVec{2} = [initMx(4) initMx(5) initMx(6)];
initVec{3} = [initMx(7) initMx(8) initMx(9)];
initVec{4} = initMx(10);
initVec{5} = initMx(11);
initVec{6} = initMx(12);
initVec{7} = initMx(13);
initVec{8} = initMx(14);

%full vector generator
for a = 1:length(initVec{1})
    for b = 1:length(initVec{2})
        for c = 1:length(initVec{3})
            vecMx{c+loopSize,a} = [initVec{1}(a) ...
                initVec{2}(b) initVec{3}(c) initVec{4:8}];
        end
        loopSize = loopSize+c;
    end
    loopSize = 0;
end

%brute force all combinations to populate appliance vector
for i = 1:size(vecMx,2)
    for j = 1:size(vecMx,1)
        for k = 1:8
            combos = nchoosek(vecMx{j,i},k);
            combosT = sum(combos,2);
            combRes{j,i}{1,k} = combosT;
        end
    end
```

```
39 end
40
41 for i = 1:size(vecMx,2)
42     for j = 1:size(vecMx,1)
43         for k = 1:8
44             combResCat(catSize+1:catSize
45                 +length(combRes{j,i}{1,k}),1) = ...
                       combRes{j,i}{1,k};
46             catSize = length(combResCat);
47         end
48     end
49 end
50
51 %apply unique index to numerical data
52 resFinal = combResCat(indStr);
53
54 %append '0' to front of combMx
55 resFinal(2:end+1,:) = resFinal(1:end,:); resFinal(1) = 0;
```

Figure C.3: `TLMCombGenerator` function.

```
%declare and populate alphabetical vectors
initVecStr{1} = ['a' 'b' 'c'];
initVecStr{2} = ['d' 'e' 'f'];
initVecStr{3} = ['g' 'h' 'i'];
initVecStr{4} = 'j';
initVecStr{5} = 'k';
initVecStr{6} = 'l';
initVecStr{7} = 'm';
initVecStr{8} = 'n';

%declare variables
loopSize = 0;
catSize = 0;

%full vector generator
for a = 1:length(initVecStr{1})
    for b = 1:length(initVecStr{2})
        for c = 1:length(initVecStr{3})
            vecMxStr{c+loopSize,a} = [initVecStr{1}(a) ...
                initVecStr{2}(b) initVecStr{3}(c) ...
                initVecStr{4:8}];
        end
        loopSize = loopSize+c;
    end
    loopSize = 0;
end

%brute force all combinations and populate alphabetical vector
for i = 1:size(vecMxStr,2)
    for j = 1:size(vecMxStr,1)
        for k = 1:8
            combosStr = nchoosek(vecMxStr{j,i},k);
            combResStr{j,i}{1,k} = cellstr(combosStr);
        end
    end
end

for i = 1:size(vecMxStr,2)
    for j = 1:size(vecMxStr,1)
```

```
38          for k = 1:8
39              combResCatStr(catSize+1:catSize
40                  +length(combResStr{j,i}{1,k}),1) = ...
                       combResStr{j,i}{1,k};
41              catSize = length(combResCatStr);
42          end
43      end
44  end
45
46  %Find unique index to apply to vectors
47  [resStrFinal, indStr, ic] = unique(combResCatStr);
48
49  %append '0' to front of combMx
50  resStrFinal(2:end+1,:) = resStrFinal(1:end,:);
51  resStrFinal(1) = cellstr('zero');
```

Figure C.4: `TLMCharGenerator` function.

```
%import individual appliance real power data
load('comboPAll.mat');

%call steady-state transformation function
ssTransformation;

%declare variables
compMx = zeros(size(comboPAll{2},1),1);
diffMx = zeros(size(comboPAll{2},2),1);

%find individual and combined steady-state differences
for i = 1:size(ssDetResP,1)
    compMx(1:size(compMx,1),i) = ...
       abs(ssDetResP(i,2)-comboPAll{2}(1:size(compMx,1),1));
end

%locate minimum differences
for i = 1:size(compMx,2)
    compRes{i,1} = indSSP(i);
    compRes{i,2} = ssDetResP(i,2);
    diffMx(i,1) = min(compMx(1:size(compMx,1),i));
    ind = find(compMx(1:size(compMx,1),i) == diffMx(i,1));
    for j = 1:length(ind);
        compRes{i,3} = ind(j);
    end

    %include match, difference and string
    compRes{i,4} = comboPAll{2}(ind);
    compRes{i,5} = diffMx(i,1);
    compRes{i,6} = comboPAll{1}(compRes{i,3},1);
end

%call string process function for compRes
strProcessorForTPM;
```

Figure C.5: `TLMLevelMatcher` function.

```
%call alphabetical combinator function
TPMCharGeneratorElimMW3LevelsQ;

endSize = 0; catSize = 0; %declare variables

%declare appliance state combination vectors
initVecQ{1} = [initMxQ(1) initMxQ(2) initMxQ(3)];
initVecQ{2} = initMxQ(4);

%full vector generator
for a = 1:length(initVecQ{1})
    vecMxQ{a,1} = [initVecQ{1}(a) initVecQ{2}];
end

%brute force all combinations to populate appliance vector
for i = 1:size(vecMxQ,1)
    for k = 1:2
        combos = nchoosek(vecMxQ{i},k);
        combosT = sum(combos,2);
        combResQ{i}{1,k} = combosT;
    end
end

for i = 1:size(vecMxQ,1)
    for k = 1:2
        combResCatQ(catSize+1:catSize
            +length(combResQ{i}{1,k}),1) = combResQ{i}{1,k};
        catSize = length(combResCatQ);
    end
end

%apply unique index to numerical data and append '0' to combMx
resFinalQ = combResCatQ(indStr);
resFinalQ(2:end+1,:) = resFinalQ(1:end,:);
```

Figure C.6: `CPMCombGenerator` function.

```
%call functions
ssTransformation; %real and reactive power ss transformation
CPMProcessorQ; %process reactive power signal
strProcessorForPandQForQ; %process strings in compResQ

%find real power matches based upon reactive power results
for i = 1:size(compResQ,1)
    switch compResQ{i,3}
        case 1
            %result contains no microwave or fridge
            CPMProcessorPNoMwNoFrg;
            cse{i,1} = 'NoMwNoFrg'; %error checking vector
        case 2
            %result must contain microwave-front, no fridge
            CPMProcessorPAllMwFNoFrg;
            cse{i,1} = 'AllMwFNoFrg';
        case 3
            %result must contain microwave-front and fridge
            CPMProcessorPAllMwFAndFrg;
            cse{i,1} = 'AllMwFAndFrg';
        case 4
            %result must contain microwave-back, no fridge
            CPMProcessorPAllMwBNoFrg;
            cse{i,1} = 'AllMwBNoFrg';
        case 5
            %result must contain microwave-back and fridge
            CPMProcessorPAllMwBAndFrg;
            cse{i,1} = 'AllMwBAndFrg';
        case 6
            %result must contain microwave-inactive, no fridge
            CPMProcessorPAllMwINoFrg;
            cse{i,1} = 'AllMwINoFrg';
        case 7
            %result must contain microwave-inactive and fridge
            CPMProcessorPAllMwIAndFrg;
            cse{i,1} = 'AllMwIAndFrg';
        case 8
            %result must contain fridge, no microwave
            CPMProcessorPAllFrgNoMw;
```

```
40              cse{i,1} = 'AllFrgNoMw';
41          otherwise
42              cse{i,1} = 'Error!';
43      end
44  end
45
46  %call function to process strings in compResP
47  strProcessorForPandQForP;
```

Figure C.7: `CPMLevelMatcher` function.

```
%set sample parameters
posLen = 250; negLen = 200; posEdgeLen = -10; negEdgeLen = 10;

%pre-allocate crop sample vectors
cropPPos = zeros(posLen,1); cropSPos = zeros(posLen,1);
cropPNeg = zeros(negLen,1); cropSNeg = zeros(negLen,1);

for x = 1:size(loadArray,1)

    load(loadArray{x,1}); %load in individual samples

    %call event detection function and find edges
    eventDetection;
    indPPos = find(edgeDetPPos > 1);
    indPNeg = find(edgeDetPNeg > 1);

    %call general or special extractor functions
    if (x == 10 || x == 11 || x == 13)
        extractorMW;
    else
        extractorGeneral;
    end
end

%average the microwave waveforms
MWWaveformAverager;

%generate the sampleArray files for correlation purposes
makeSampleCellArrayCorrelation;

%save sample parameters for further processing
lenInfo = zeros(4,1);
lenInfo(1) = posLen; lenInfo(2) = negLen;
lenInfo(3) = posEdgeLen; lenInfo(4) = negEdgeLen;
```

Figure C.8: `EECExtractorIndividual` function.

```
%extract positive edge samples
for i = 1:length(indPPos)
    for n = 1:posLen
        cropPPos(n,i) = P(indPPos(i)+n-1+posEdgeLen);
    end
end

%extract negative edge samples
for i = 1:length(indPNeg)
    for n = 1:negLen %capture negative edges
        cropPNeg(n,i) = P(indPNeg(i)+n-(negLen-1)+negEdgeLen);
    end
end

%create variable name cell array
varCell{1,1} = [saveNames{x} 'PPos'];
varCell{2,1} = [saveNames{x} 'PNeg'];

%rename samples
eval([varCell{1,1},'= cropPPos;']);
eval([varCell{2,1},'= cropPNeg;']);

%save varCell variables to file
excluder = ['^' saveNames{x}];
save(saveArray{x,1},varCell{:,1},'-regexp',excluder);
```

Figure C.9: `EECExtractorGeneral` function.

```
%pre-allocate sample vectors
cropPPos1 = zeros(posLen,1);
cropPPos2 = zeros(posLen,1);
cropPNeg = zeros(negLen,1);

%call event detection function and find edges
eventDetection;
indPPos = find(edgeDetPPos > 1);
indPNeg = find(edgeDetPNeg > 1);

%extract front 1 edge samples
count = 1;
for i = 1:2:length(indPPos)
    for n = 1:posLen %capture positive edges
        cropPPos1(n,count) = P(indPPos(i)+n-1+posEdgeLen);
    end
    count=count+1;
end

%extract front 2 edge samples
count = 1;
for i = 2:2:length(indPPos)
    for n = 1:posLen %capture positive edges
        cropPPos2(n,count) = P(indPPos(i)+n-1+posEdgeLen);
    end
    count=count+1;
end

%extract negative edge samples
for i = 1:length(indPNeg)
    for n = 1:negLen %capture negative edges
        cropPNeg(n,i) = P(indPNeg(i)+n-(negLen-1)+negEdgeLen);
    end
end

%create variable name cell array:
varCell{1,1} = [saveNames{x} 'PPos1'];
varCell{2,1} = [saveNames{x} 'PPos2'];
varCell{3,1} = [saveNames{x} 'PNeg'];
```

```
40
41 %rename samples
42 eval([varCell{1,1},'= cropPPos1;']);
43 eval([varCell{2,1},'= cropPPos2;']);
44 eval([varCell{3,1},'= cropPNeg;']);
45
46 %save varCell variables to file:
47 excluder = ['^' saveNames{x}];
48 save(saveArray{x,1},varCell{:,1},'-regexp',excluder);
```

Figure C.10: `EECExtractorMicrowave` function.

```
%average front 1 edge samples
tempMean11 = mean(mwLPPos1,2);
tempMean12 = mean(mwMPPos1,2);
tempMean13 = mean(mwHPPos1,2);
tempMx1 = [tempMean11 tempMean12 tempMean13];
meanPPos1 = mean(tempMx1,2);

%average front 2 edge samples
tempMean21 = mean(mwLPPos2,2);
tempMean22 = mean(mwMPPos2,2);
tempMean23 = mean(mwHPPos2,2);
tempMx2 = [tempMean21 tempMean22 tempMean23];
meanPPos2 = mean(tempMx2,2);

meanPPos = [meanPPos1 meanPPos2]; %collate front edges

%average negative edge samples
tempMean31 = mean(mwLPNeg,2); tempMean32 = mean(mwMPNeg,2);
tempMean33 = mean(mwHPNeg,2);
tempMx3 = [tempMean31 tempMean32 tempMean33];
meanPNeg = mean(tempMx3,2);

%create variable name cell array:
varCell{1,1} = ['mwA' 'PPos'];
varCell{2,1} = ['mwA' 'PNeg'];

%rename samples
eval([varCell{1,1},'= meanPPos;']);
eval([varCell{2,1},'= meanPNeg;']);

%save varCell variables to file
saveFile = ['C:\****\****\****' saveFolder '\' 'mwA' '.mat'];
excluder = ['^' 'mwA'];
save(saveFile,varCell{:,1},'-regexp',excluder);
```

Figure C.11: `EECMWSampleAverager` function.

```
%call event detection function and find edges
indPPos = find(edgeDetPPos > 1);
indPNeg = find(edgeDetPNeg > 1);

%preallocate vectors for extracted edge samples
samplePos = zeros(posLen,1);
sampleNeg = zeros(negLen,1);

%pre-allocate matrix for edge correlation values
combP = zeros(size(sampleArrayP,1)+2,length(P));

%extract and correlate positive event edges
for i = 1:length(indPPos)
    c = 1;
    %extract samples from aggregated signal
    for n = 1:posLen
        samplePos(n) = P(indPPos(i)+n-1+posEdgeLen);
    end

    %correlate samples against library and store values
    for a = 1:size(sampleArrayP,1)
        for b = 1:size(sampleArrayP{a,1},2)
            corrArrayPPos{1,i}(c,1) = ...
                corr(samplePos,sampleArrayP{a,1}(:,b));
            combP(1,indPPos(i)) = indPPos(i);
            c = c+1;
        end
    end
    combP(2:end,indPPos(i)) = corrArrayPPos{1,i};
end

%extract and correlate negative event edges
for i = 1:length(indPNeg)
    c = 1;
    %extract samples from aggregated signal
    for n = 1:negLen
        sampleNeg(n) = P(indPNeg(i)+n-(negLen-1)+negEdgeLen);
    end

```

```
39      %correlate samples against library and store values
40      for a = 1:size(sampleArrayP,1)
41          for b = 1:size(sampleArrayP{a,1},2)
42              corrArrayPNeg{1,i}(c,1) = ...
                    corr(sampleNeg,sampleArrayP{a,2}(:,1));
43              combP(1,indPNeg(i)) = -indPNeg(i);
44              c = c+1;
45          end
46      end
47      combP(2:end,indPNeg(i)) = corrArrayPNeg{1,i};
48      combP(8,indPNeg(i)) = 0;
49  end
50
51  %eliminate zero rows in combP matrix
52  combSlim = find(combP(1,:)~=0);
53  combP = combP(:,combSlim);
```

Figure C.12: `EECCorrelator` function.

Appendix D

DETAILED TLM RESULTS

This appendix contains tables of the results produced when TLM is applied to the combined measurements, as referenced from Chapter 5. These results are reported at the measurement-by-measurement level. TLM provided the lowest overall accuracy of the three disaggregation techniques, as discussed in the overview of results contained in Section 5.5.

TLM is a steady-state technique that considers only the real power component of the measurements that it processes, as discussed in Section 1.3.1. As the least sophisticated of the disaggregation techniques, it might be expected to offer the worst accuracy. The relative performance of TLM, CPM and EEC is discussed comprehensively in Chapter 5 and Appendix G.

A legend to the appliance name abbreviations utilised in the tables contained within this appendix is provided below.

Appliance	**Abbreviation**
Toaster	tst
Kettle	ktl
Lamp	lmp
Heater-Low	ht1
Heater-Medium	ht2
Heater-High	ht3
Snackwich	snw
Fan-Low	fnL
Fan-Medium	fnM
Fan-High	fnH
Microwave-Front	mwF
Microwave-Back	mwB
Microwave-Inactive	mwI
Fridge	frg

Table D.1: TLM results for Combined Measurement 1.

Time (s)	**Measured SS (W)**	**Matched SS (W)**	**Output**	**Ground Truth Data**	**Score**
9.9	752.62	753.62	fnL,ht1,mwF	tst	0
19.9	765.45	765.37	fnH,ht1,mwF	tst,fnL	0
29.9	731.57	731.31	tst	tst	1
39.9	762.18	760.95	fnM,tst	tst,fnM	1
50.04	728.64	728.76	fnM,ht1,frg,lmp	tst	0
59.92	766.65	767.49	fnM,ht1,mwI,frg-lmp	tst,fnH	0
70.7	725.77	724.92	fnL,ht1,frg,lmp	tst	0
226.82	0.00	0.00	null	null	1
				Total (%)	**38**

Table D.2: TLM results for Combined Measurement 2.

Time (s)	Measured SS (W)	Matched SS (W)	Output	Ground Truth Data	Score
8.92	201.73	202.93	fnL,frg,lmp	mwF	0
10.84	1024.72	1024.81	fnM,mwF,lmp,tst	mwB	0
16.56	39.84	38.73	mwI	mwI	1
19.02	776.92	777.09	fnM,lmp,snw	ts,mwI	0
41.32	931.96	932.77	fnH,mwF,snw	ts,mwF	0
43.02	1748.70	1749.48	fnM,ht2,mwI,frg,lmp,tst	ts,mwB	0
46.74	766.41	765.68	fnH,mwI,snw	ts,mwI	0
59.28	727.74	727.81	ht1,mwF	ts	0
79.02	764.79	765.37	fnH,ht1,mwF	ts,mwI	0
91.08	931.06	931.05	fnL,mwI,frg,lmp,snw	ts,mwF	0
92.92	1746.32	1745.64	fnL,ht2,mwI,frg,lmp,tst	ts,mwB	0
96.3	769.18	768.87	fnH,tst	ts,mwI	0
99.3	36.46	37.56	fnH	mwI	0
121.08	202.52	202.93	fnL,frg,lmp	mwF	0
122.78	1023.96	1024.81	fnM,mwF,lmp,tst	mwB	0
126.48	37.21	37.56	fnH	mwI	0
139.26	0.00	0.00	null	null	1
				Total %	**12**

Table D.3: TLM results for Combined Measurement 3.

Time (s)	Measured SS (W)	Matched SS (W)	Output	Ground Truth Data	Score
9	37.11	37.56	fnH	mwI	0
19	784.95	785.00	fnH,lmp,snw	tst,mwI	0
32	934.25	934.23	fnL,frg,lmp,tst	tst,mwF	0
34	1812.59	1812.77	fnL,ht2,mwF,frg,snw	tst,mwB	0
38	769.66	770.03	mwI,tst	tst,mwI	1
50	730.14	731.31	tst	tst	1
59	775.83	775.40	fnH,ht1,mwI,frg,lmp	tst,mwI	0
61	933.56	934.23	fnL,frg,lmp,tst	tst,mwF	0
63	1842.69	1841.19	fnH,mwF,frg,lmp,snw,tst	tst,mwB	0
77	766.46	765.68	fnH,mwI,snw	tst,mwI	0
89	727.53	727.81	ht1,mwF	tst	0
109	764.20	763.65	fnL,ht1,mwI,frg,lmp	tst,mwI	0
122	918.89	918.75	fnM,mwI,frg,tst	tst,mwF	0
123	1775.88	1775.23	fnM,mwF,frg,snw,tst	tst,mwB	0
127	759.36	757.77	fnM,mwI,snw	tst,mwI	0
129	36.63	37.56	fnH	mwI	0
140	0.17	0.00	null	null	1
149	41.70	38.73	mwI	mwI	1
150	201.17	202.93	fnL,frg,lmp	mwF	0
152	1088.76	1089.11	fnL,mwB	mwB	0
166	50.19	58.05	lmp	mwI	0
169	778.86	777.09	fnM,lmp,snw	tst,mwI	0
180	924.45	924.85	fnM,mwF,snw	tst,mwF	0
181	1809.74	1810.74	mwB,lmp,snw	tst,mwB	0
196	765.48	765.37	fnH,ht1,mwF	tst,mwI	0
199	37.13	37.56	fnH	mwI	0
210	204.04	202.93	fnL,frg,lmp	mwF	0
211	1056.27	1056.19	mwF,frg,tst	mwB	0
219	0.00	0.00	null	null	1
				Total %	**17**

Table D.4: TLM results for Combined Measurement 2.

Time (s)	Measured SS (W)	Matched SS (W)	Output	Ground Truth Data	Score
9	698.23	699.11	ht1,frg,lmp	snw	0
29	753.25	753.62	fnL,ht1,mwF	snw,lmp	0
49	778.23	777.09	fnM,lmp,snw	snw,lmp,fnL	0
69	745.94	747.44	lmp,snw	snw,lmp	1
89	774.68	775.40	fnH,ht1,mwI,frg,lmp	snw,lmp,fnM	0
110	740.42	737.84	ht1,mwI,frg,lmp	snw,lmp	0
130	776.55	777.09	fnM,lmp,snw	snw,lmp,fnH	0
144	95.88	95.61	fnH,lmp	lmp,fnH	1
149	58.00	58.05	lmp	lmp	1
190	0.11	0.00	null	null	1
210	57.75	58.05	lmp	lmp	1
230	88.93	87.69	fnM,lmp	lmp,fnM	1
271	30.56	29.65	fnM	fnM	1
290	0.00	0.00	null	null	1
				Total %	**57**

Table D.5: TLM results for Combined Measurement 5.

Time (s)	**Measured SS (W)**	**Matched SS (W)**	**Output**	**Ground Truth Data**	**Score**
10	522.18	522.00	ht1	ht1	1
30	579.80	580.05	ht1,lmp	ht1,lmp	1
50	607.69	605.86	fnL,ht1,lmp	ht1,lmp,fnL	1
70	579.25	580.05	ht1,lmp	ht1,lmp	1
90	610.76	609.69	fnM,ht1,lmp	ht1,lmp,fnM	1
110	579.14	580.05	ht1,lmp	ht1,lmp	1
130	617.06	617.61	fnH,ht1,lmp	ht1,lmp,fnH	1
150	578.56	580.05	ht1,lmp	ht1,lmp	1
170	57.39	58.05	lmp	lmp	1
190	0.08	0.00	null	null	1
210	57.54	58.05	lmp	lmp	1
230	89.06	87.69	fnM,lmp	lmp,fnM	1
250	610.01	609.69	fnM,ht1,lmp	lmp,fnM,ht1	1
270	552.25	551.64	fnM,ht1	fnM,ht1	1
290	521.19	522.00	ht1	ht1	1
310	0.00	0.00	null	null	1
				Total %	**1.00**

Table D.6: TLM results for Combined Measurement 6.

Time (s)	Measured SS (W)	Matched SS (W)	Output	Ground Truth Data	Score
10	145.31	144.88	fnL,frg	frg	0
59	911.60	908.42	frg,lmp,tst	frg,ht2	0
80	1099.36	1098.13	fnL,mwF,frg,lmp,snw	frg,ht2,mwF	0
81	2010.42	2010.10	fnM,ht3,snw	frg,ht2,mwB	0
94	946.57	947.15	mwI,frg,lmp,tst	frg,ht2,mwI	0
100	172.23	177.11	frg,lmp	frg,mwI	0
107	320.23	324.88	mwF,frg	frg,mwF	1
108	1217.89	1219.93	fnH,mwB,frg	frg,mwB	0
119	2547.29	2547.91	fnM,ht2,mwF,frg,snw,tst	ht3,frg,mwB	0
124	1464.63	1464.79	fnH,ht1,mwI,frg,lmp,snw	frg,ht3,mwI	0
137	1604.28	1604.30	fnL,mwI,frg,snw,tst	frg,ht3,mwF	0
138	2262.76	2261.76	fnM,ht2,mwI,snw,tst	frg,ht3,mwB	0
140	1423.19	1425.39	fnH,ht3,mwI,lmp	frg,ht3	0
159	129.41	126.42	fnM,mwI,lmp	frg	0
179	338.10	350.69	fnL,mwF,frg	frg,mwF	0
181	1206.00	1208.18	fnL,mwB,frg	frg,mwB	0
193	167.34	157.79	mwI,frg	frg,mwI	1
199	37.53	37.56	fnH	mwI	0
218	182.35	183.60	fnL,mwI,frg	mwF	0
219	1737.91	1739.45	fnM,ht2,mwF,tst	ht2,mwB	0
223	815.94	815.81	fnM,mwI,lmp,snw	ht2,mwI	0
239	981.25	982.90	fnM,mwF,lmp,snw	frg,ht2,mwI	0
248	1107.37	1109.88	fnH,mwF,frg,lmp,snw	frg,ht2,mwF	0
250	1942.25	1942.70	ht1,snw,tst	frg,ht2,mwB	0
253	945.84	945.98	fnH,frg,lmp,tst	frg,ht2,mwI	0
259	168.49	177.11	frg,lmp	frg,mwI	0
278	315.37	324.88	mwF,frg	frg,mwF	1
279	371.93	362.44	fnH,mwF,frg	mwF	0
280	1072.39	1072.32	mwF,frg,lmp,snw	mwB	0
283	38.30	38.73	mwI	mwI	1
299	0.00	0.00	null	null	1
				Total %	**16**

Table D.7: TLM results for Combined Measurement 7.

Time (s)	**Measured SS (W)**	**Matched SS (W)**	**Output**	**Ground Truth Data**	**Score**
10	58.19	58.05	lmp	lmp	1
29	580.82	580.05	ht1,lmp	lmp,ht1	1
50	1326.13	1328.62	fnH,ht3	lmp,ht1,tst	0
70	790.33	789.35	lmp,tst	lmp,tst	1
90	2079.13	2078.67	fnH,mwI,ktl	lmp,ht2,tst	0
110	787.71	786.17	mwI,lmp,snw	lmp,tst	0
129	727.51	727.81	ht1,mwF	tst	0
135	1.55	0.00	null	null	1
149	1293.92	1295.25	fnL,ht1,lmp,snw	ht3	0
169	1350.75	1350.08	ht1,mwI,lmp,tst	ht3,lmp	0
190	2085.08	2086.24	fnL,ktl,lmp	ht3,lmp,tst	0
210	787.52	786.17	mwI,lmp,snw	lmp,tst	0
229	58.42	58.05	lmp	lmp	1
249	0.00	0.00	null	null	1
				Total %	**43**

Table D.8: TLM results for Combined Measurement 8.

Time (s)	**Measured SS (W)**	**Matched SS (W)**	**Output**	**Ground Truth Data**	**Score**
9	163.40	157.79	mwI,frg	frg	0
39	2161.60	2160.18	mwI,frg,ktl	frg,ktl	0
49	2212.47	2212.07	fnL,ht3,mwF,snw	frg,ktl,lmp	0
59	2151.52	2151.10	fnM,frg,ktl	frg,ktl	0
64	2207.54	2208.20	mwF,ktl	frg,ktl,lmp	0
70	2147.15	2147.26	fnL,frg,ktl	frg,ktl	0
79	149.70	148.71	fnM,frg	frg	0
84	2142.02	2141.43	ht3,frg,tst	frg,ktl	0
89	2201.74	2199.48	ht3,frg,lmp,tst	frg,ktl,lmp	0
95	2141.62	2141.43	ht3,frg,tst	frg,ktl	0
100	2003.83	2002.39	ktl	ktl	1
				Total %	**9**

Table D.9: TLM results for Combined Measurement 9.

Time (s)	Measured SS (W)	Matched SS (W)	Output	Ground Truth Data	Score
10	745.91	747.44	lmp,snw	tst	0
30	936.25	937.12	mwF,tst	tst,mwF	1
32	1769.86	1771.39	fnL,mwF,frg,snw,tst	tst,mwB	0
40	2567.62	2567.86	fnL,mwB,lmp,snw,tst	tst,mwB,ht2	0
50	1504.96	1504.90	fnM,ht1,mwF,lmp,snw	tst,ht2	0
60	728.58	728.76	fnM,ht1,frg,lmp	tst	0
80	929.55	929.31	fnH,ht2,frg	tst,mwF	0
82	1757.74	1757.39	fnH,ht2,mwI,frg,lmp,tst	tst,mwB	0
90	1828.08	1828.87	ht2,mwF,frg,tst	tst,mwB,lmp	0
95	1769.04	1767.85	ht2,mwF,lmp,tst	tst,mwB	0
100	1825.79	1824.52	fnH,ht2,mwF,frg,snw	tst,mwB,lmp	0
105	1764.60	1763.50	fnH,ht2,mwF,lmp,snw	tst,mwB	0
110	729.51	728.76	fnM,ht1,frg,lmp	tst	0
120	2018.09	2018.01	fnH,ht3,snw	tst,ht3	0
140	1293.33	1292.03	ht1,mwI,tst	ht3	0
160	1494.91	1494.96	fnL,ht1,mwI,frg,lmp,tst	ht3,mwF	0
162	2284.73	2286.22	ht3,mwF,lmp,tst	ht3,mwB	0
167	1330.27	1330.46	ht1,frg,snw	ht3,mwI	0
170	39.27	38.73	mwI	mwI	1
180	211.54	214.68	fnH,frg,lmp	mwF	0
181	989.96	990.82	fnH,mwF,lmp,snw	mwB	0
190	1527.03	1526.62	fnL,ht2,mwI,snw	mwB,ht1	0
197	561.47	560.73	ht1,mwI	mwI,ht1	1
210	527.88	522.00	ht1	ht1	1
230	730.92	731.31	tst	ht1,mwF	0
232	1508.17	1508.39	fnM,lmp,snw,tst	ht1,mwB	0
240	1578.54	1578.49	mwI,frg,snw,tst	ht1,mwB,lmp	0
247	618.79	618.77	ht1,mwI,lmp	ht1,mwI,lmp	1
260	558.87	559.56	fnH,ht1	ht1,mwI	0
270	521.87	522.00	ht1	ht1	1
280	0.00	0.00	null	null	1
				Total %	**23**

Appendix E

DETAILED CPM RESULTS

This appendix contains tables of the results produced when CPM is applied to the combined measurements, as referenced from Chapter 5. These results are reported at the measurement-by-measurement level. CPM provided the second best overall accuracy of the three disaggregation techniques, improving on the performance of TLM, as discussed in the overview of results contained in Section 5.5.

CPM is a steady-state technique that considers both the real and reactive power components of the measurements that it processes, as discussed in Section 1.3.1. Due to the additional layer of information provided by the inclusion of reactive power data, it might be expected to improve upon the accuracy of TLM. The relative performance of TLM, CPM and EEC is discussed comprehensively in Chapter 5 and Appendix G.

A legend to the appliance name abbreviations utilised in the tables contained within this appendix is provided below.

Appliance	**Abbreviation**
Toaster	tst
Kettle	ktl
Lamp	lmp
Heater-Low	ht1
Heater-Medium	ht2
Heater-High	ht3
Snackwich	snw
Fan-Low	fnL
Fan-Medium	fnM
Fan-High	fnH
Microwave-Front	mwF
Microwave-Back	mwB
Microwave-Inactive	mwI
Fridge	frg

Table E.1: CPM results for Combined Measurement 1.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
10	-0.47	0.00	null	752.62	757.12	fnL,tst	tst	0
20	0.00	0.00	null	765.45	768.87	fnH,tst	tst,fnL	0
30	0.00	0.00	null	731.57	731.31	tst	tst	1
40	0.00	0.00	null	762.18	760.95	fnM,tst	tst,fnM	1
50	0.00	0.00	null	728.64	726.95	fnH,snw	tst	0
60	0.00	0.00	null	766.65	768.87	fnH,tst	tst,fnH	1
71	0.04	0.00	null	725.77	726.95	fnH,snw	tst	0
227	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**50**

Table E.2: CPM results for Combined Measurement 2.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
9	658.03	654.34	mwF	201.73	205.81	mwF	mwF	1
11	363.74	384.41	mwB	1024.72	1063.30	mwB	mwB	1
17	37.02	31.39	mwI	39.84	38.73	mwI	mwI	1
19	0.79	0.00	null	776.92	777.09	fnM,lmp,snw	tst,mwI	0
41	621.90	654.34	mwF	931.96	932.77	fnH,mwF,snw	tst,mwF	0
43	358.74	384.41	mwB	1748.70	1752.69	mwB,snw	tst,mwB	0
47	-0.78	0.00	null	766.41	768.87	fnH,tst	tst,mwI	0
59	0.00	0.00	null	727.74	726.95	fnH,snw	tst	0
79	1.02	0.00	null	764.79	760.95	fnM,tst	tst,mwI	0
91	631.12	654.34	mwF	931.06	932.77	fnH,mwF,snw	tst,mwF	0
93	366.66	384.41	mwB	1746.32	1752.69	mwB,snw	tst,mwB	0
96	9.35	0.00	null	769.18	768.87	fnH,tst	tst,mwI	0
99	31.62	31.39	mwI	36.46	38.73	mwI	mwI	1
121	642.46	654.34	mwF	202.52	205.81	mwF	mwF	1
123	374.01	384.41	mwB	1023.96	1063.30	mwB	mwB	1
126	32.00	31.39	mwI	37.21	38.73	mwI	mwI	1
139	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**47**

Table E.3: CPM results for Combined Measurement 3.

Time (s)	Measured SS (VAR)	Matched SS (VAR)	Reactive Power Output	Measured SS (W)	Matched SS (W)	Real Power Output	Ground Truth Data	Score
9	33.29	31.39	mwI	37.11	38.73	mwI	mwI	1
19	0.59	0.00	null	784.95	785.00	fnH,lmp,snw	tst,mwI	0
32	619.64	654.34	mwF	934.25	932.77	fnH,mwF,snw	tst,mwF	0
34	420.21	384.41	mwB	1812.59	1810.74	mwB,lmp,snw	tst,mwB	0
38	-0.92	0.00	null	769.66	768.87	fnH,tst	tst,mwI	0
50	0.08	0.00	null	730.14	731.31	tst	tst	1
59	7.83	0.00	null	775.83	777.09	fnM,lmp,snw	tst,mwI	0
61	624.25	654.34	mwF	933.56	932.77	fnH,mwF,snw	tst,mwF	0
63	412.05	384.41	mwB	1842.69	1840.39	fnM,mwB,lmp,snw	tst,mwB	0
77	1.65	0.00	null	766.46	768.87	fnH,tst	tst,mwI	0
89	0.00	0.00	null	727.53	726.95	fnH,snw	tst	0
109	0.57	0.00	null	764.20	760.95	fnM,tst	tst,mwI	0
122	620.79	654.34	mwF	918.89	921.02	fnL,mwF,snw	tst,mwF	0
123	413.32	384.41	mwB	1775.88	1778.50	fnL,mwB,snw	tst,mwB	0
127	5.45	0.00	null	759.36	760.95	fnM,tst	tst,mwI	0
129	31.97	31.39	mwI	36.63	38.73	mwI	mwI	1
140	0.21	0.00	null	0.17	0.00	null	null	1
149	37.34	31.39	mwI	41.70	38.73	mwI	mwI	1
150	653.27	654.34	mwF	201.17	205.81	mwF	mwF	1
152	401.80	384.41	mwB	1088.76	1089.11	fnL,mwB	mwB	0
166	38.04	31.39	mwI	50.19	38.73	mwI	mwI	1
169	1.13	0.00	null	778.86	777.09	fnM,lmp,snw	tst,mwI	0
180	624.87	654.34	mwF	924.45	924.85	fnM,mwF,snw	tst,mwF	0
181	383.38	384.41	mwB	1809.74	1810.74	mwB,lmp,snw	tst,mwB	0
196	-2.43	0.00	null	765.48	768.87	fnH,tst	tst,mwI	0
199	32.54	31.39	mwI	37.13	38.73	mwI	mwI	1
210	649.33	654.34	mwF	204.04	205.81	mwF	mwF	1
211	384.30	384.41	mwB	1056.27	1063.30	mwB	mwB	1
219	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**38**

Table E.4: CPM results for Combined Measurement 4.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
9	-0.13	0.00	null	698.23	689.39	snw	snw	1
29	0.11	0.00	null	753.25	757.12	fnL,tst	snw,lmp	0
49	0.00	0.00	null	778.23	777.09	fnM,lmp,snw	snw,lmp,fnL	0
69	0.00	0.00	null	745.94	747.44	lmp,snw	snw,lmp	1
89	0.00	0.00	null	774.68	773.25	fnL,lmp,snw	snw,lmp,fnM	0
110	0.00	0.00	null	740.42	747.44	lmp,snw	snw,lmp	1
130	0.33	0.00	null	776.55	777.09	fnM,lmp,snw	snw,lmp,fnH	0
144	0.94	0.00	null	95.88	95.61	fnH,lmp	lmp,fnH	1
149	0.01	0.00	null	58.00	58.05	lmp	lmp	1
190	-0.08	0.00	null	0.11	0.00	null	null	1
210	-0.06	0.00	null	57.75	58.05	lmp	lmp	1
230	0.01	0.00	null	88.93	87.69	fnM,lmp	lmp,fnM	1
271	0.03	0.00	null	30.56	29.65	fnM	fnM	1
290	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**71**

Table E.5: CPM results for Combined Measurement 5.

Time (s)	Measured SS (VAR)	Matched SS (VAR)	Reactive Power Output	Measured SS (W)	Matched SS (W)	Real Power Output	Ground Truth Data	Score
10	0.13	0.00	null	522.18	522.00	ht1	ht1	1
30	-0.12	0.00	null	579.80	580.05	ht1,lmp	ht1,lmp	1
50	0.00	0.00	null	607.69	605.86	fnL,ht1,lmp	ht1,lmp,fnL	1
70	0.00	0.00	null	579.25	580.05	ht1,lmp	ht1,lmp	1
90	0.00	0.00	null	610.76	609.69	fnM,ht1,lmp	ht1,lmp,fnM	1
110	0.00	0.00	null	579.14	580.05	ht1,lmp	ht1,lmp	1
130	0.00	0.00	null	617.06	617.61	fnH,ht1,lmp	ht1,lmp,fnH	1
150	0.15	0.00	null	578.56	580.05	ht1,lmp	ht1,lmp	1
170	0.13	0.00	null	57.39	58.05	lmp	lmp	1
190	-0.13	0.00	null	0.08	0.00	null	null	1
210	-0.10	0.00	null	57.54	58.05	lmp	lmp	1
230	0.22	0.00	null	89.06	87.69	fnM,lmp	lmp,fnM	1
250	0.21	0.00	null	610.01	609.69	fnM,ht1,lmp	lmp,fnM,ht1	1
270	0.00	0.00	null	552.25	551.64	fnM,ht1	fnM,ht1	1
290	0.17	0.00	null	521.19	522.00	ht1	ht1	1
310	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**1.00**

Table E.6: CPM results for Combined Measurement 6.

Time (s)	Measured SS (VAR)	Matched SS (VAR)	Reactive Power Output	Measured SS (W)	Matched SS (W)	Real Power Output	Ground Truth Data	Score
10	158.15	155.69	frg	145.31	144.88	fnL,frg	frg	0
59	139.93	155.69	frg	911.60	908.42	frg,lmp,tst	frg,ht2	0
80	762.89	810.04	mwF,frg	1099.36	1098.13	fnL,mwF,frg,lmp,snw	frg,ht2,mwF	0
81	465.33	540.10	mwB,frg	2010.42	2009.28	fnH,mwB,frg,lmp,tst	frg,ht2,mwB	0
94	179.29	187.08	mwI,frg	946.57	947.15	mwI,frg,lmp,tst	frg,ht2,mwI	0
100	193.16	187.08	mwI,frg	172.23	183.60	fnL,mwI,frg	frg,mwI	0
107	798.33	810.04	mwF,frg	320.23	324.88	mwF,frg	frg,mwF	1
108	480.27	540.10	mwB,frg	1217.89	1219.93	fnH,mwB,frg	frg,mwB	0
119	432.25	384.41	mwB	2547.29	2551.19	fnL,ht2,mwB,snw	ht3,frg,mwB	0
124	166.76	155.69	frg	1464.63	1467.98	fnH,ht1,frg,lmp,tst	frg,ht3,mwI	0
137	745.45	810.04	mwF,frg	1604.28	1604.00	fnL,ht1,mwF,frg,tst	frg,ht3,mwF	0
138	553.74	540.10	mwB,frg	2262.76	2393.76	ht1,mwB,frg,snw	frg,ht3,mwB	0
140	127.72	155.69	frg	1423.19	1426.07	fnH,ht1,frg,lmp,snw	frg,ht3	0
159	155.44	155.69	frg	129.41	119.07	frg	frg	1
179	796.23	810.04	mwF,frg	338.10	350.69	fnL,mwF,frg	frg,mwF	0
181	446.20	384.41	mwB	1206.00	1158.91	fnH,mwB,lmp	frg,mwB	0
193	192.19	187.08	mwI,frg	167.34	157.79	mwI,frg	frg,mwI	1
199	32.65	31.39	mwI	37.53	38.73	mwI	mwI	1
218	669.53	654.34	mwF	182.35	205.81	mwF	mwF	1
219	415.96	384.41	mwB	1737.91	1752.69	mwB,snw	ht2,mwB	0
223	2.46	0.00	null	815.94	815.16	fnL,lmp,tst	ht2,mwI	0
239	191.02	187.08	mwI,frg	981.25	984.71	fnH,mwI,frg,lmp,tst	frg,ht2,mwI	0
248	760.55	810.04	mwF,frg	1107.37	1109.88	fnH,mwF,frg,lmp,snw	frg,ht2,mwF	0
250	431.31	384.41	mwB	1942.25	1931.59	fnH,ht2,mwB,lmp	frg,ht2,mwB	0
253	179.27	187.08	mwI,frg	945.84	947.15	mwI,frg,lmp,tst	frg,ht2,mwI	0
259	190.82	187.08	mwI,frg	168.49	157.79	mwI,frg	frg,mwI	1
278	776.24	810.04	mwF,frg	315.37	324.88	mwF,frg	frg,mwF	1
279	632.91	654.34	mwF	371.93	301.42	fnH,mwF,lmp	mwF	0
280	377.66	384.41	mwB	1072.39	1063.30	mwB	mwB	1
283	32.55	31.39	mwI	38.30	38.73	mwI	mwI	1
299	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**32**

Table E.7: CPM results for Combined Measurement 7.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
10	0.27	0.00	null	58.19	58.05	lmp	lmp	1
29	0.42	0.00	null	580.82	580.05	ht1,lmp	lmp,ht1	1
50	0.55	0.00	null	1326.13	1328.62	fnH,ht3	lmp,ht1,tst	0
70	0.37	0.00	null	790.33	789.35	lmp,tst	lmp,tst	1
90	0.21	0.00	null	2079.13	2080.41	ht3,lmp,tst	lmp,ht2,tst	0
110	0.49	0.00	null	787.71	789.35	lmp,tst	lmp,tst	1
129	0.87	0.00	null	727.51	726.95	fnH,snw	tst	0
135	-0.18	0.00	null	1.55	0.00	null	null	1
149	-0.21	0.00	null	1293.92	1295.25	fnL,ht1,lmp,snw	ht3	0
169	-0.42	0.00	null	1350.75	1349.10	ht3,lmp	ht3,lmp	1
190	0.09	0.00	null	2085.08	2086.24	fnL,ktl,lmp	ht3,lmp,tst	0
210	0.69	0.00	null	787.52	789.35	lmp,tst	lmp,tst	1
229	0.21	0.00	null	58.42	58.05	lmp	lmp	1
249	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**64**

Table E.8: CPM results for Combined Measurement 8.

Time (s)	**Measured SS (VAR)**	**Matched SS (VAR)**	**Reactive Power Output**	**Measured SS (W)**	**Matched SS (W)**	**Real Power Output**	**Ground Truth Data**	**Score**
9	163.40	155.69	frg	163.40	156.63	fnH,frg	frg	0
39	108.67	155.69	frg	2161.60	2159.01	fnH,frg,ktl	frg,ktl	0
49	104.68	155.69	frg	2212.47	2209.15	fnM,frg,ktl,lmp	frg,ktl,lmp	0
59	107.61	155.69	frg	2151.52	2151.10	fnM,frg,ktl	frg,ktl	0
64	104.33	155.69	frg	2207.54	2209.15	fnM,frg,ktl,lmp	frg,ktl,lmp	0
70	107.09	155.69	frg	2147.15	2147.26	fnL,frg,ktl	frg,ktl	0
79	162.50	155.69	frg	149.70	148.71	fnM,frg	frg	0
84	108.67	155.69	frg	2142.02	2141.43	ht3,frg,tst	frg,ktl	0
89	103.84	155.69	frg	2201.74	2199.48	ht3,frg,lmp,tst	frg,ktl,lmp	0
95	106.29	155.69	frg	2141.62	2141.43	ht3,frg,tst	frg,ktl	0
100	0.00	0.00	null	2003.83	2002.39	ktl	ktl	1
							Total (%)	**9**

Table E.9: CPM results for Combined Measurement 9.

Time (s)	Measured SS (VAR)	Matched SS (VAR)	Reactive Power Output	Measured SS (W)	Matched SS (W)	Real Power Output	Ground Truth Data	Score
10	0.84	0.00	null	745.91	747.44	lmp,snw	tst	0
30	642.10	654.34	mwF	936.25	937.12	mwF,tst	tst,mwF	1
32	354.16	384.41	mwB	1769.86	1778.50	fnL,mwB,snw	tst,mwB	0
40	333.97	384.41	mwB	2567.62	2567.86	fnL,mwB,lmp,snw,tst	tst,mwB,ht2	0
50	0.83	0.00	null	1504.96	1504.56	fnL,lmp,snw,tst	tst,ht2	0
60	0.62	0.00	null	728.58	726.95	fnH,snw	tst	0
80	635.02	654.34	mwF	929.55	932.77	fnH,mwF,snw	tst,mwF	0
82	344.53	384.41	mwB	1757.74	1752.69	mwB,snw	tst,mwB	0
90	333.97	384.41	mwB	1828.08	1824.25	fnM,mwB,tst	tst,mwB,lmp	0
95	330.93	384.41	mwB	1769.04	1778.50	fnL,mwB,snw	tst,mwB	0
100	329.91	384.41	mwB	1825.79	1824.25	fnM,mwB,tst	tst,mwB,lmp	0
105	329.04	384.41	mwB	1764.60	1752.69	mwB,snw	tst,mwB	0
110	-1.26	0.00	null	729.51	731.31	tst	tst	1
120	-0.35	0.00	null	2018.09	2018.01	fnH,ht3,snw	tst,ht3	0
140	0.25	0.00	null	1293.33	1295.25	fnL,ht1,lmp,snw	ht3	0
160	622.61	654.34	mwF	1494.91	1496.68	fnH,ht1,mwF,tst	ht3,mwF	0
162	318.33	384.41	mwB	2284.73	2274.69	ht1,mwB,snw	ht3,mwB	0
167	2.88	0.00	null	1330.27	1328.62	fnH,ht3	ht3,mwI	0
170	34.44	31.39	mwI	39.27	38.73	mwI	mwI	1
180	671.14	654.34	mwF	211.54	205.81	mwF	mwF	1
181	331.38	384.41	mwB	989.96	1063.30	mwB	mwB	1
190	310.72	384.41	mwB	1527.03	1585.30	ht1,mwB	mwB,ht1	1
197	0.32	0.00	null	561.47	559.56	fnH,ht1	mwI,ht1	0
210	17.68	31.39	mwI	527.88	560.73	ht1,mwI	ht1	0
230	652.09	654.34	mwF	730.92	727.81	ht1,mwF	ht1,mwF	1
232	318.97	384.41	mwB	1508.17	1585.30	ht1,mwB	ht1,mwB	1
240	307.81	384.41	mwB	1578.54	1585.30	ht1,mwB	ht1,mwB,lmp	0
247	1.91	0.00	null	618.79	617.61	fnH,ht1,lmp	ht1,mwI,lmp	0
260	0.00	0.00	null	558.87	559.56	fnH,ht1	ht1,mwI	0
270	-0.39	0.00	null	521.87	522.00	ht1	ht1	1
280	0.00	0.00	null	0.00	0.00	null	null	1
							Total (%)	**32**

Appendix F

DETAILED EEC RESULTS

This appendix contains tables of the results produced when EEC is applied to the combined measurements, as referenced from Chapter 5. These results are reported at the measurement-by-measurement level. EEC provided the best overall accuracy of the three disaggregation techniques, as discussed in the overview of results contained in Section 5.5.

EEC is a transient technique that considers and compares real power waveform event edges found within the measurements that it processes, as discussed in Section 1.3.1. Due to the presence of unique identifying features within appliance state change waveform event edges, EEC might be expected to provide the best accuracy. The relative performance of TLM, CPM and EEC is discussed comprehensively in Chapter 5 and Appendix G.

A legend to the appliance name abbreviations utilised in the tables contained within this appendix is provided below. In each of the tables of results, positive or negative signs indicate the direction of the edge at each point in time where an event is detected. The EEC implementation ensures that individual appliance and combined measurement edge samples of opposing directions are never correlated. Thus it may be assumed that a positive column denotes the correlation value of a positive appliance edge sample, and a negative col-

umn a negative edge sample. Note the special case of the microwave, where the negative edge of a microwave magnatron event corresponds to a negative edge for the second front section, the first section having no negative edge associated with it.

Appliance	**Abbreviation**
Fan-High	fnH
Fan-Medium	fnM
Lamp	fnL
Heater-Low	ht3
Heater-Medium	ht2
Heater-High	ht1
Microwave-Front1	mwF1
Microwave-Front2	mwF2
Microwave-Inactive	mwI
Fridge	frg
Kettle	ktl
Lamp	lmp
Snackwich	snw
Toaster	tst

Table F.1: EEC results for Combined Measurement 1.

Appliance	**+10s**	**+20s**	**-30s**	**+40s**	**-50s**	**+60s**	**-71s**	**-227s**	**Total (%)**
fnH	0.8357	0.5967	0.9444	0.7251	0.9160	0.9561	0.9663	0.9945	
fnM	0.9610	0.8015	0.9695	0.9925	0.9358	0.8234	0.9851	0.9948	
fnL	0.8617	0.9527	0.9704	0.8748	0.9522	0.5833	0.9912	0.9928	
ht3	0.1210	-0.0463	0.7992	0.1172	0.7812	0.1030	0.8273	0.8861	
ht2	0.1241	-0.0436	0.8101	0.1201	0.7913	0.1059	0.8378	0.8952	
ht1	0.1110	-0.0452	0.9664	0.1115	0.9345	0.0880	0.9845	0.9982	
mwF1	-0.8946	-0.8301	0.0000	-0.9707	0.0000	-0.6922	0.0000	0.0000	
mwF2	0.0947	0.4597	0.3441	0.1812	0.4206	-0.1595	0.4656	0.5080	
mwI	0.4166	0.3853	0.9184	0.3465	0.8921	0.4965	0.9410	0.9777	
frg	0.0295	0.1104	0.9656	-0.0045	0.9356	0.0387	0.9842	0.9940	
ktl	0.1829	0.0091	0.6872	0.1753	0.6755	0.1576	0.7171	0.7878	
lmp	-0.0580	0.0775	0.9229	-0.0741	0.8966	-0.0361	0.9465	0.9822	
snw	0.1104	-0.0561	0.9670	0.1079	0.9387	0.0934	0.9861	0.9984	
tst	0.9896	0.8348	0.9648	0.9507	0.9309	0.8757	0.9827	0.9989	
Ground Truth	**tst**	**fnL**	**fnL**	**fnM**	**fnM**	**fnH**	**fnH**	**tst**	
Hard Score	1	1	1	1	0	1	0	1	**75**
Soft Score	1	1	1	1	1	1	1	1	**100**

Table F.2: EEC results for Combined Measurement 2.

Appliance	**+9s**	**+11s**	**-17s**	**+19s**	**+41s**	**+43s**	**-47s**	**-59s**	**+79s**
fnH	-0.6113	-0.5623	0.9904	0.2510	-0.6775	0.1381	0.3030	0.9925	0.9061
fnM	-0.9157	-0.2590	0.9977	0.6039	-0.9517	0.3865	0.2807	0.9728	0.7690
fnL	-0.8365	-0.1423	0.9966	0.5588	-0.8046	0.7257	0.2799	0.9719	0.5439
ht3	0.0603	0.3715	0.8697	0.8784	0.0351	0.0586	0.3092	0.9208	-0.0348
ht2	0.0574	0.3695	0.8794	0.8799	0.0327	0.0595	0.3088	0.9281	-0.0320
ht1	0.0578	0.3923	0.9996	0.8674	0.0410	0.0700	0.2870	0.9801	-0.0452
mwF1	0.9863	0.2652	0.0000	-0.3776	0.9888	-0.4215	0.0000	0.0000	-0.6802
mwF2	-0.2229	0.5773	0.4966	0.2063	-0.1630	0.8337	0.9181	0.5026	-0.1813
mwI	-0.3971	-0.6306	0.9700	-0.4518	-0.4171	0.0307	0.3085	0.9865	0.5865
frg	-0.0989	-0.5475	0.9969	-0.6336	-0.0819	-0.0739	0.2887	0.9720	0.2395
ktl	0.0072	0.3377	0.7663	0.9066	-0.0191	0.0857	0.2963	0.8388	0.0182
lmp	-0.0830	-0.4792	0.9752	-0.8330	-0.0623	-0.0722	0.3084	0.9886	0.1234
snw	0.0686	0.3774	0.9998	0.8733	0.0436	0.0526	0.2871	0.9814	-0.0440
tst	-0.8277	-0.3463	0.9997	0.6881	-0.8573	0.4417	0.2885	0.9815	0.8035
Ground Truth	**mwF1**	**mwF2**	**mwF2**	**tst**	**mwF1**	**mwF2**	**mwF2**	**mwI**	**mwI**
Hard Score	1	1	0	0	1	1	1	0	0
Soft Score	1	1	0	0	1	1	1	0	0

Appliance	+91s	+93s	-96s	-99s	+121s	+123s	-126s	-139s	Total (%)
fnH	-0.6274	0.2079	0.0361	0.2049	-0.6851	0.1558	0.2923	0.9720	
fnM	-0.9427	0.4793	0.0282	0.2102	-0.9327	0.3997	0.2681	0.9927	
fnL	-0.8363	0.8056	0.0287	0.2092	-0.7830	0.7343	0.2673	0.9975	
ht3	-0.0043	0.0460	0.0422	0.1792	0.1037	0.0563	0.3030	0.8290	
ht2	-0.0071	0.0471	0.0401	0.1830	0.1013	0.0571	0.3023	0.8397	
ht1	-0.0026	0.0591	0.0318	0.2073	0.1100	0.0681	0.2747	0.9918	
mwF1	0.9890	-0.5196	0.0000	0.0000	0.9853	-0.4335	0.0000	0.0000	
mwF2	-0.2115	0.7841	0.8576	-0.5901	-0.1522	0.8297	0.9117	0.4869	
mwI	-0.3612	0.0810	0.0354	0.2047	-0.4641	0.0428	0.2989	0.9452	
frg	-0.0473	-0.0564	0.0374	0.2016	-0.1163	-0.0723	0.2759	0.9925	
ktl	-0.0579	0.0780	0.0425	0.1581	0.0500	0.0843	0.2926	0.7167	
lmp	-0.0207	-0.0578	0.0380	0.2013	-0.1240	-0.0693	0.2984	0.9514	
snw	0.0042	0.0396	0.0310	0.2085	0.1120	0.0501	0.2749	0.9929	
tst	-0.8555	0.5259	0.0328	0.2064	-0.8303	0.4566	0.2763	0.9901	
Ground Truth	**mwF1**	**mwF2**	**mwF2**	**tst**	**mwF1**	**mwF2**	**mwF2**	**mwI**	
Hard Score	1	0	1	0	1	1	1	0	**59**
Soft Score	1	0	1	0	1	1	1	0	**59**

Table F.3: EEC results for Combined Measurement 3.

Appliance	**+9s**	**+19s**	**+32s**	**+34s**	**-38s**	**-50s**	**+59s**	**+61s**	**+63s**	**-77s**
fnH	0.5428	0.1844	-0.6120	0.1519	0.2201	0.9629	-0.4306	-0.6962	-0.6015	0.9297
fnM	0.5803	0.5435	-0.9043	0.4208	0.2004	0.9854	-0.6470	-0.9567	-0.3064	0.8355
fnL	0.4162	0.5072	-0.8262	0.7576	0.2005	0.9910	-0.8960	-0.7859	-0.1795	0.8343
ht3	0.4349	0.9127	0.0969	0.0556	0.2283	0.8181	0.0491	0.0117	0.3276	0.9999
ht2	0.4371	0.9140	0.0941	0.0565	0.2271	0.8289	0.0474	0.0095	0.3256	0.9997
ht1	0.4277	0.9020	0.0943	0.0726	0.2067	0.9838	0.0439	0.0195	0.3521	0.8581
mwF1	-0.3860	-0.3115	0.9850	-0.4610	0.0000	0.0000	0.6855	0.9739	0.2993	0.0000
mwF2	-0.1518	0.2095	-0.2204	0.8483	0.9425	0.4602	-0.6146	-0.1415	0.5754	0.5033
mwI	0.1092	-0.5153	-0.4207	0.0404	0.2237	0.9359	-0.2876	-0.4118	-0.6196	0.9634
frg	0.1002	-0.6701	-0.1181	-0.0724	0.2096	0.9844	-0.0893	-0.0673	-0.4956	0.8349
ktl	0.4714	0.9365	0.0442	0.0837	0.2204	0.7055	0.0035	-0.0425	0.2914	0.9833
lmp	-0.2804	-0.8712	-0.1159	-0.0710	0.2247	0.9418	-0.0679	-0.0417	-0.4330	0.9575
snw	0.4283	0.9083	0.1052	0.0494	0.2066	0.9852	0.0578	0.0201	0.3339	0.8620
tst	0.6664	0.6300	-0.8124	0.4682	0.2083	0.9819	-0.6966	-0.8642	-0.3994	0.8641
Ground Truth	**mwI**	**tst**	**mwF1**	**mwF2**	**mwF2**	**mwI**	**mwI**	**mwF1**	**mwF2**	**mwF2**
Hard Score	0	0	1	1	1	0	0	1	1	0
Soft Score	0	0	1	1	1	0	1	1	1	0

Appliance	-89s	+109s	+122s	+123s	-127s	-129s	-140s	+149s	+150s	+152s
fnH	0.9921	0.8679	-0.6507	0.0347	0.3690	0.2068	0.9991	-0.4547	-0.6981	-0.6150
fnM	0.9737	0.5399	-0.9347	0.3014	0.3694	0.2155	0.9808	-0.6745	-0.9436	-0.3195
fnL	0.9716	0.3539	-0.8186	0.6561	0.3685	0.2120	0.9794	-0.9136	-0.7797	-0.1915
ht3	0.9183	-0.5673	0.0646	0.0694	0.3268	0.1803	0.9240	0.0543	0.0731	0.3253
ht2	0.9255	-0.5652	0.0619	0.0703	0.3291	0.1839	0.9314	0.0525	0.0708	0.3233
ht1	0.9808	-0.5749	0.0667	0.0816	0.3713	0.2090	0.9881	0.0518	0.0808	0.3481
mwF1	0.0000	-0.5897	0.9925	-0.3433	0.0000	0.0000	0.0000	0.7129	0.9799	0.3111
mwF2	0.4996	-0.2528	-0.1865	0.8447	0.9239	-0.3540	0.5146	-0.5722	-0.1401	0.5744
mwI	0.9856	0.9147	-0.4237	-0.0402	0.3616	0.2089	0.9919	-0.3052	-0.4536	-0.6267
frg	0.9722	0.5750	-0.1088	-0.0947	0.3771	0.2019	0.9801	-0.0957	-0.1075	-0.4935
ktl	0.8352	-0.5223	0.0106	0.0903	0.2893	0.1583	0.8401	0.0071	0.0190	0.2883
lmp	0.9878	0.6185	-0.0908	-0.0917	0.3640	0.2016	0.9945	-0.0738	-0.0991	-0.4285
snw	0.9818	-0.5751	0.0731	0.0642	0.3705	0.2106	0.9891	0.0628	0.0815	0.3318
tst	0.9826	0.5309	-0.8413	0.3445	0.3715	0.2077	0.9897	-0.7189	-0.8449	-0.4142
Ground Truth	**mwI**	**mwI**	**mwF1**	**mwF2**	**mwF2**	**tst**	**mwI**	**mwI**	**mwF1**	**mwF2**
Hard Score	0	1	1	1	1	0	0	0	1	1
Soft Score	0	1	1	1	1	0	0	1	1	1

Appliance	-166s	+169s	+180s	+181s	-196s	-199s	+210s	+211s	-219s	Total (%)
fnH	0.8501	0.8047	-0.7339	-0.7036	0.9742	0.2160	-0.7409	-0.5852	0.8516	
fnM	0.7256	0.9568	-0.9461	-0.3877	0.9081	0.2243	-0.9460	-0.2839	0.7274	
fnL	0.7244	0.8442	-0.7605	-0.2337	0.9067	0.2203	-0.7559	-0.1607	0.7264	
ht3	0.9843	0.2495	0.0784	0.2872	0.9886	0.1890	0.0790	0.3448	0.9849	
ht2	0.9806	0.2526	0.0763	0.2847	0.9914	0.1922	0.0769	0.3429	0.9813	
ht1	0.7541	0.2377	0.0877	0.3050	0.9250	0.2177	0.0884	0.3671	0.7559	
mwF1	0.0000	-0.8544	0.9686	0.3587	0.0000	0.0000	0.9647	0.2818	0.0000	
mwF2	0.4657	0.0860	-0.1085	0.5696	0.5151	-0.2989	-0.0988	0.5805	0.4706	
mwI	0.9012	0.3112	-0.4778	-0.6504	0.9926	0.2179	-0.4851	-0.6236	0.9024	
frg	0.7249	-0.0675	-0.1148	-0.4354	0.9072	0.2114	-0.1278	-0.5132	0.7269	
ktl	0.9998	0.3096	0.0237	0.2459	0.9452	0.1668	0.0242	0.3097	0.9999	
lmp	0.8916	-0.1823	-0.1067	-0.3862	0.9897	0.2111	-0.1112	-0.4505	0.8929	
snw	0.7591	0.2393	0.0869	0.2940	0.9279	0.2191	0.0875	0.3508	0.7610	
tst	0.7618	0.9978	-0.8515	-0.4892	0.9294	0.2168	-0.8533	-0.3757	0.7636	
Ground Truth	**mwF2**	**tst**	**mwF1**	**mwF2**	**mwF2**	**tst**	**mwF1**	**mwF2**	**mwF2**	
Hard Score	0	1	1	1	0	0	1	1	0	**55**
Soft Score	0	1	1	1	1	0	1	1	0	**66**

Table F.4: EEC results for Combined Measurement 4.

Appliance	**+9s**	**+29s**	**+49s**	**-69s**	**+89s**	**-110s**	**+130s**	**-144s**
fnH	-0.2173	0.3658	0.5845	0.9379	0.6946	0.9687	0.9866	0.9938
fnM	0.1394	0.1109	0.7862	0.9579	0.9909	0.9897	0.8078	0.9954
fnL	0.1490	0.0781	0.9379	0.9699	0.8932	0.9921	0.5646	0.9946
ht3	0.9981	-0.6807	-0.0937	0.7999	0.1936	0.8262	-0.0443	0.8832
ht2	0.9979	-0.6792	-0.0911	0.8104	0.1965	0.8369	-0.0413	0.8924
ht1	0.9911	-0.6749	-0.0960	0.9569	0.1883	0.9884	-0.0582	0.9985
mwF1	0.0945	-0.2314	-0.8273	0.0000	-0.9539	0.0000	-0.7119	0.0000
mwF2	0.1996	-0.2644	0.4462	0.4605	0.2232	0.4356	-0.1863	0.5040
mwI	-0.8155	0.7771	0.4120	0.9128	0.2756	0.9424	0.6156	0.9765
frg	-0.8009	0.9682	0.1541	0.9576	-0.0699	0.9882	0.1533	0.9946
ktl	0.9938	-0.6599	-0.0390	0.6915	0.2510	0.7143	0.0111	0.7839
lmp	-0.9845	0.7988	0.1258	0.9181	-0.1515	0.9482	0.1082	0.9810
snw	0.9984	-0.6838	-0.1033	0.9598	0.1845	0.9890	-0.0538	0.9990
tst	0.2235	0.0929	0.8118	0.9541	0.9599	0.9870	0.8451	0.9989
Ground Truth	**snw**	**lmp**	**fnL**	**fnL**	**fnM**	**fnM**	**fnH**	**snw**
Hard Score	1	0	1	1	1	0	1	1
Soft Score	1	0	1	1	1	1	1	1

Appliance	-149s	-190s	+210s	+230s	-271s	-290s	Total (%)
fnH	0.9974	0.9791	0.2370	0.7719	0.9569	0.9486	
fnM	0.9907	0.9999	-0.1257	0.9601	0.8798	0.9689	
fnL	0.9892	0.9984	-0.1429	0.8437	0.8786	0.9786	
ht3	0.9011	0.8351	-0.9710	-0.0906	0.9933	0.8090	
ht2	0.9095	0.8459	-0.9706	-0.0877	0.9949	0.8195	
ht1	0.9956	0.9991	-0.9635	-0.0951	0.8990	0.9679	
mwF1	0.0000	0.0000	-0.0891	-0.9858	0.0000	0.0000	
mwF2	0.5085	0.4930	-0.2558	0.1450	0.4973	0.4753	
mwI	0.9841	0.9519	0.8380	0.5159	0.9814	0.9227	
frg	0.9900	0.9992	0.9124	0.1565	0.8793	0.9692	
ktl	0.8079	0.7219	-0.9646	-0.0318	0.9604	0.6995	
lmp	0.9878	0.9584	0.9971	0.1291	0.9773	0.9285	
snw	0.9962	0.9986	-0.9718	-0.1000	0.9023	0.9703	
tst	0.9966	0.9984	-0.1935	0.9026	0.9040	0.9655	
Ground Truth	**fnH**	**lmp**	**lmp**	**fnM**	**lmp**	**fnM**	
Hard Score	1	0	1	1	0	0	**64**
Soft Score	1	0	1	1	0	1	**79**

Table F.5: EEC results for Combined Measurement 5.

Appliance	**+10s**	**+30s**	**+50s**	**-70s**	**+90s**	**-110s**	**+130s**	**-150s**	**-170s**
fnH	-0.1920	0.2638	0.5171	0.9459	0.7869	0.9737	0.9623	0.9577	0.9918
fnM	0.1653	0.1166	0.7968	0.9657	0.9633	0.9925	0.8192	0.9782	0.9967
fnL	0.1708	0.0705	0.9736	0.9773	0.8228	0.9910	0.5781	0.9872	0.9963
ht3	0.9998	-0.3071	0.0648	0.8067	-0.0853	0.8336	0.0825	0.8168	0.8751
ht2	0.9998	-0.3057	0.0672	0.8171	-0.0825	0.8441	0.0855	0.8274	0.8846
ht1	0.9927	-0.3025	0.0669	0.9650	-0.0909	0.9921	0.0680	0.9772	0.9992
mwF1	0.0703	-0.1419	-0.8125	0.0000	-0.9824	0.0000	-0.6915	0.0000	0.0000
mwF2	0.1974	-0.2994	0.5331	0.4534	0.1148	0.4754	-0.1706	0.4634	0.5046
mwI	-0.8001	0.4923	0.2602	0.9200	0.5219	0.9472	0.5146	0.9317	0.9726
frg	-0.7923	0.8143	0.0146	0.9655	0.1560	0.9912	0.0549	0.9781	0.9961
ktl	0.9974	-0.2891	0.1176	0.6974	-0.0266	0.7223	0.1374	0.7061	0.7733
lmp	-0.9825	0.4735	-0.0369	0.9257	0.1260	0.9538	-0.0154	0.9375	0.9775
snw	0.9998	-0.3099	0.0553	0.9679	-0.0946	0.9917	0.0729	0.9794	0.9996
tst	0.2510	0.1445	0.8324	0.9623	0.9034	0.9917	0.8719	0.9749	0.9993
Ground Truth	**ht1**	**lmp**	**fnL**	**fnL**	**fnM**	**fnM**	**fnH**	**fnH**	**ht1**
Hard Score	0	0	1	1	1	1	1	0	0
Soft Score	1	0	1	1	1	1	1	1	0

Appliance	-190s	+210s	+230s	+250s	-270s	-290s	-310s	Total (%)
fnH	0.9988	0.2629	0.7763	-0.1903	0.9948	0.9799	0.9923	
fnM	0.9878	0.1230	0.9721	0.1657	0.9871	0.9902	0.9966	
fnL	0.9862	0.0663	0.8479	0.1699	0.9859	0.9895	0.9954	
ht3	0.9105	-0.1849	-0.0475	0.9999	0.9007	0.8549	0.8769	
ht2	0.9185	-0.1835	-0.0446	0.9999	0.9091	0.8648	0.8863	
ht1	0.9935	-0.1822	-0.0530	0.9923	0.9923	0.9914	0.9992	
mwF1	0.0000	-0.1057	-0.9835	0.0710	0.0000	0.0000	0.0000	
mwF2	0.5141	-0.3122	0.1417	0.1956	0.5014	0.4977	0.5018	
mwI	0.9879	0.4133	0.4895	-0.7991	0.9821	0.9583	0.9735	
frg	0.9869	0.7419	0.1234	-0.7922	0.9862	0.9896	0.9959	
ktl	0.8208	-0.1672	0.0116	0.9975	0.8086	0.7502	0.7756	
lmp	0.9911	0.3618	0.0878	-0.9824	0.9857	0.9636	0.9783	
snw	0.9943	-0.1879	-0.0569	0.9999	0.9928	0.9915	0.9995	
tst	0.9947	0.1713	0.9191	0.2514	0.9932	0.9913	0.9995	
Ground Truth	**lmp**	**lmp**	**fnM**	**ht1**	**lmp**	**fnM**	**ht1**	
Hard Score	0	0	1	0	0	0	0	**38**
Soft Score	0	0	1	1	0	0	0	**56**

Table F.6: EEC results for Combined Measurement 6.

Appliance	**+10s**	**+59s**	**+80s**	**+81s**	**-94s**	**-100s**	**+107s**	**+108s**
fnH	0.5659	-0.1834	-0.5863	-0.7991	0.9074	0.9048	-0.7546	-0.8049
fnM	0.3607	0.1752	-0.8961	-0.4799	0.8030	0.7995	-0.9499	-0.4857
fnL	0.2344	0.1822	-0.8478	-0.2999	0.8018	0.7984	-0.7475	-0.3056
ht3	-0.0801	0.9999	0.0686	0.2383	0.9984	0.9980	0.0418	0.2377
ht2	-0.0781	0.9999	0.0655	0.2355	0.9970	0.9966	0.0397	0.2350
ht1	-0.1071	0.9921	0.0617	0.2513	0.8277	0.8243	0.0539	0.2515
mwF1	-0.2964	0.0611	0.9754	0.4305	0.0000	0.0000	0.9485	0.4355
mwF2	-0.5351	0.2019	-0.2496	0.5063	0.4997	0.4956	-0.0878	0.5001
mwI	0.4178	-0.7955	-0.3880	-0.6701	0.9467	0.9449	-0.4680	-0.6725
frg	0.2589	-0.7928	-0.1014	-0.3646	0.8026	0.7989	-0.1033	-0.3636
ktl	-0.0435	0.9981	0.0160	0.1926	0.9920	0.9927	-0.0132	0.1917
lmp	0.1789	-0.9824	-0.0886	-0.3276	0.9397	0.9376	-0.0769	-0.3270
snw	-0.0868	0.9997	0.0766	0.2459	0.8319	0.8287	0.0503	0.2455
tst	0.4768	0.2616	-0.8146	-0.5784	0.8342	0.8309	-0.8651	-0.5838
Ground Truth	**frg**	**ht2**	**mwF1**	**mwF2**	**mwF2**	**ht2**	**mwF1**	**mwF2**
Hard Score	0	1	1	1	0	0	1	1
Soft Score	0	1	1	1	0	1	1	1

Appliance	+119s	-124s	+137s	+138s	-140s	-159s	+179s	+181s
fnH	0.1386	0.9999	-0.1343	0.4981	0.0578	0.9882	-0.6455	-0.5889
fnM	0.2952	0.9784	0.0137	0.7841	0.0439	0.9366	-0.9204	-0.2921
fnL	0.5380	0.9769	0.4171	0.6250	0.0449	0.9352	-0.8057	-0.1692
ht3	0.1322	0.9305	-0.0167	0.0006	0.0634	0.9748	0.1002	0.3473
ht2	0.1329	0.9376	-0.0149	0.0014	0.0631	0.9791	0.0976	0.3452
ht1	0.1194	0.9863	0.0030	-0.0150	0.0520	0.9507	0.1024	0.3660
mwF1	-0.2811	0.0000	-0.0755	-0.7840	0.0000	0.0000	0.9903	0.2907
mwF2	0.4392	0.5120	0.4238	0.2665	0.3541	0.5199	-0.1943	0.5777
mwI	-0.0144	0.9945	-0.0631	0.2724	0.0485	0.9988	-0.4352	-0.6282
frg	-0.1041	0.9775	0.0094	0.0070	0.0589	0.9359	-0.0932	-0.5181
ktl	0.1560	0.8490	-0.0076	0.0389	0.0632	0.9185	0.0472	0.3115
lmp	-0.1274	0.9966	0.0054	0.0004	0.0586	0.9976	-0.1129	-0.4513
snw	0.1255	0.9874	-0.0168	-0.0053	0.0498	0.9530	0.1085	0.3534
tst	0.3752	0.9881	0.1057	0.6502	0.0544	0.9542	-0.8198	-0.3831
Ground Truth	ht3	mwF2	mwF1	mwF2	mwF2	ht3	mwF1	mwF2
Hard Score	0	0	0	0	1	0	1	1
Soft Score	0	0	1	0	1	0	1	1

Appliance	-193s	-199s	+218s	+219s	-223s	+239s	+248s	+250s
fnH	0.9005	0.9509	-0.7983	-0.3473	-0.0286	0.5255	-0.6745	0.1489
fnM	0.7935	0.8694	-0.9500	-0.0030	-0.0233	0.3237	-0.9516	0.4228
fnL	0.7923	0.8678	-0.7535	0.3714	-0.0235	0.2097	-0.8092	0.7617
ht3	0.9974	0.9960	0.0807	0.1907	-0.0370	-0.1203	0.0323	0.0526
ht2	0.9957	0.9971	0.0782	0.1911	-0.0380	-0.1183	0.0298	0.0536
ht1	0.8186	0.8894	0.0932	0.2002	-0.0236	-0.1387	0.0384	0.0705
mwF1	0.0000	0.0000	0.9497	-0.0363	0.0000	-0.2730	0.9895	-0.4651
mwF2	0.4908	0.5002	-0.0706	0.7279	0.7908	-0.5257	-0.1662	0.8487
mwI	0.9416	0.9779	-0.5206	-0.3446	-0.0323	0.4258	-0.4144	0.0404
frg	0.7928	0.8682	-0.1420	-0.2133	-0.0141	0.3095	-0.0832	-0.0678
ktl	0.9939	0.9672	0.0244	0.1900	-0.0394	-0.0858	-0.0220	0.0808
lmp	0.9341	0.9732	-0.1183	-0.2256	-0.0308	0.2190	-0.0604	-0.0684
snw	0.8230	0.8928	0.0895	0.1895	-0.0248	-0.1268	0.0408	0.0465
tst	0.8253	0.8947	-0.8744	0.0268	-0.0239	0.4321	-0.8585	0.4688
Ground Truth	**mwF2**	**frg**	**mwF1**	**mwF2**	**mwF2**	**frg**	**mwF1**	**mwF2**
Hard Score	0	0	1	1	1	0	1	1
Soft Score	0	0	1	1	1	0	1	1

Appliance	-253s	-259s	+278s	-279s	+280s	-283s	-299s	Total (%)
fnH	0.0889	1.0000	-0.6012	0.0467	0.3500	0.1456	0.9824	
fnM	0.0663	0.9783	-0.9053	0.0436	0.5544	0.1262	0.9412	
fnL	0.0666	0.9769	-0.7852	0.0447	0.8471	0.1258	0.9387	
ht3	0.1165	0.9308	0.0983	0.0393	-0.0239	0.1632	0.9511	
ht2	0.1137	0.9379	0.0961	0.0401	-0.0224	0.1615	0.9563	
ht1	0.0729	0.9862	0.1031	0.0479	-0.0166	0.1316	0.9529	
mwF1	0.0000	0.0000	0.9795	0.0000	-0.5904	0.0000	0.0000	
mwF2	0.8929	0.5162	-0.2018	0.2177	0.6302	0.8788	0.5081	
mwI	0.0957	0.9946	-0.4056	0.0390	0.2240	0.1519	0.9878	
frg	0.0754	0.9775	-0.0801	0.0542	0.0637	0.1353	0.9401	
ktl	0.1236	0.8495	0.0474	0.0338	0.0167	0.1633	0.8871	
lmp	0.0964	0.9967	-0.1066	0.0445	0.0397	0.1515	0.9875	
snw	0.0728	0.9874	0.1062	0.0456	-0.0318	0.1314	0.9547	
tst	0.0747	0.9880	-0.7905	0.0495	0.6223	0.1329	0.9559	
Ground Truth	**mwF2**	**ht2**	**mwF1**	**frg**	**mwF2**	**mwF2**	**mwI**	
Hard Score	1	0	1	0	0	1	1	**55**
Soft Score	1	0	1	0	0	1	1	**61**

Table F.7: EEC results for Combined Measurement 7.

Appliance	**+10s**	**+29s**	**+50s**	**-70s**	**+90s**	**-110s**	**-129s**	**-135s**
fnH	0.1486	-0.1886	0.1889	0.9997	-0.2265	0.8821	0.9890	0.9016
fnM	0.1839	0.1688	0.5519	0.9753	0.1274	0.7680	0.9947	0.7949
fnL	0.1385	0.1743	0.5192	0.9738	0.1453	0.7669	0.9934	0.7937
ht3	0.2845	0.9999	0.9086	0.9356	0.9420	0.9936	0.8713	0.9976
ht2	0.2858	0.9999	0.9099	0.9425	0.9416	0.9912	0.8808	0.9960
ht1	0.2858	0.9923	0.8982	0.9837	0.9325	0.7945	0.9969	0.8200
mwF1	-0.0586	0.0677	-0.3221	0.0000	0.0785	0.0000	0.0000	0.0000
mwF2	-0.1781	0.1990	0.2193	0.5084	0.2534	0.4851	0.4888	0.4937
mwI	0.0213	-0.7983	-0.5099	0.9959	-0.8273	0.9271	0.9695	0.9424
frg	0.3581	-0.7926	-0.6671	0.9743	-0.9455	0.7674	0.9938	0.7943
ktl	0.2992	0.9977	0.9330	0.8565	0.9360	0.9976	0.7692	0.9937
lmp	-0.1058	-0.9826	-0.8674	0.9976	-0.9838	0.9188	0.9744	0.9350
snw	0.2818	0.9999	0.9041	0.9850	0.9426	0.7992	0.9971	0.8243
tst	0.2655	0.2545	0.6377	0.9857	0.1927	0.8016	0.9971	0.8266
Ground Truth	**lmp**	**ht1**	**tst**	**ht1**	**ht2**	**ht2**	**lmp**	**tst**
Hard Score	0	0	0	0	0	0	0	0
Soft Score	0	1	0	0	0	0	0	0

Appliance	+149s	+169s	+190s	-210s	-229s	-249s	Total (%)
fnH	-0.2295	0.1391	0.5993	0.9858	0.9010	0.9632	
fnM	0.1203	0.2459	0.8705	0.9312	0.7941	0.8886	
fnL	0.1400	0.2123	0.7859	0.9298	0.7929	0.8872	
ht3	0.9247	0.4224	0.5808	0.9781	0.9975	0.9942	
ht2	0.9243	0.4237	0.5834	0.9821	0.9958	0.9961	
ht1	0.9151	0.4224	0.5685	0.9458	0.8192	0.9073	
mwF1	0.0793	-0.0951	-0.7006	0.0000	0.0000	0.0000	
mwF2	0.2596	-0.1149	0.1638	0.5190	0.4938	0.5211	
mwI	-0.8235	-0.0876	-0.0398	0.9980	0.9419	0.9862	
frg	-0.9586	0.2101	-0.3559	0.9304	0.7935	0.8881	
ktl	0.9183	0.4391	0.6306	0.9244	0.9938	0.9587	
lmp	-0.9747	-0.2512	-0.5199	0.9965	0.9345	0.9825	
snw	0.9254	0.4193	0.5722	0.9482	0.8236	0.9105	
tst	0.1822	0.3373	0.9352	0.9495	0.8259	0.9122	
Ground Truth	**ht3**	**lmp**	**tst**	**ht3**	**tst**	**lmp**	
Hard Score	0	0	1	0	0	0	**7**
Soft Score	0	0	1	0	0	0	**14**

Table F.8: EEC results for Combined Measurement 8.

Appliance	**+9s**	**+39s**	**+49s**	**-59s**	**+64s**	**-70s**	**-79s**	**+84s**
fnH	0.5306	-0.1463	0.1947	0.9789	0.3367	0.9822	0.8479	-0.1868
fnM	0.3308	0.2151	0.3267	0.9502	0.2228	0.9552	0.7225	0.1818
fnL	0.2119	0.2172	0.2953	0.9485	0.2150	0.9529	0.7214	0.1918
ht3	-0.0852	0.9989	0.4762	0.9256	-0.1673	0.9260	0.9837	0.9965
ht2	-0.0831	0.9990	0.4779	0.9320	-0.1654	0.9323	0.9799	0.9965
ht1	-0.1075	0.9907	0.4723	0.9591	-0.1615	0.9637	0.7512	0.9894
mwF1	-0.2700	0.0228	-0.1661	0.0000	-0.2060	0.0000	0.0000	0.0493
mwF2	-0.5342	0.2034	-0.0781	0.4229	-0.1844	0.4985	0.4736	0.2292
mwI	0.4028	-0.7723	-0.0992	0.9785	0.4319	0.9813	0.8994	-0.8028
frg	0.2722	-0.7853	0.1394	0.9484	0.6787	0.9536	0.7219	-0.8223
ktl	-0.0503	0.9998	0.4975	0.8524	-0.1426	0.8511	1.0000	0.9946
lmp	0.1838	-0.9785	-0.3098	0.9790	0.3298	0.9817	0.8897	-0.9892
snw	-0.0916	0.9983	0.4722	0.9607	-0.1720	0.9650	0.7563	0.9964
tst	0.4452	0.3022	0.4259	0.9612	0.2821	0.9659	0.7590	0.2618
Ground Truth	**frg**	**ktl**	**lmp**	**lmp**	**lmp**	**lmp**	**ktl**	**ktl**
Hard Score	0	1	0	1	0	0	1	0
Soft Score	0	1	0	1	0	0	1	0

Appliance	+89s	-95s	-100s	Total (%)
fnH	0.4614	0.9855	0.9776	
fnM	0.1897	0.9625	0.9985	
fnL	0.1444	0.9611	0.9965	
ht3	-0.7600	0.9216	0.8338	
ht2	-0.7582	0.9283	0.8446	
ht1	-0.7570	0.9702	0.9975	
mwF1	-0.3338	0.0000	0.0000	
mwF2	-0.2521	0.4444	0.4871	
mwI	0.8613	0.9821	0.9505	
frg	0.9287	0.9603	0.9979	
ktl	-0.7334	0.8430	0.7207	
lmp	0.8508	0.9833	0.9569	
snw	-0.7638	0.9718	0.9970	
tst	0.1603	0.9721	0.9970	
Ground Truth	**lmp**	**lmp**	**frg**	
Hard Score	0	0	0	**27**
Soft Score	0	0	0	**27**

Table F.9: EEC results for Combined Measurement 9.

Appliance	**+10s**	**+30s**	**+32s**	**+40s**	**-50s**	**-60s**	**+80s**	**+82s**
fnH	0.2835	-0.5912	-0.5768	-0.1770	0.9128	0.9881	-0.6718	-0.5556
fnM	0.6352	-0.9077	-0.2678	0.1834	0.8107	0.9987	-0.9384	-0.2561
fnL	0.5913	-0.8509	-0.1502	0.1920	0.8096	0.9972	-0.8006	-0.1441
ht3	0.8573	0.0469	0.3784	0.9993	0.9990	0.8617	0.0767	0.3761
ht2	0.8589	0.0439	0.3764	0.9993	0.9979	0.8716	0.0742	0.3742
ht1	0.8464	0.0413	0.3981	0.9918	0.8349	0.9999	0.0809	0.3964
mwF1	-0.4135	0.9807	0.2733	0.0534	0.0000	0.0000	0.9903	0.2644
mwF2	0.2132	-0.2469	0.5820	0.2027	0.4988	0.4978	-0.1637	0.5717
mwI	-0.4182	-0.3756	-0.6467	-0.7905	0.9508	0.9661	-0.4430	-0.6326
frg	-0.6138	-0.0834	-0.5650	-0.7880	0.8102	0.9979	-0.1156	-0.5651
ktl	0.8878	-0.0059	0.3438	0.9982	0.9902	0.7558	0.0227	0.3425
lmp	-0.8108	-0.0675	-0.4883	-0.9805	0.9440	0.9715	-0.1028	-0.4869
snw	0.8518	0.0551	0.3845	0.9991	0.8391	0.9999	0.0851	0.3821
tst	0.7184	-0.8256	-0.3572	0.2707	0.8414	0.9999	-0.8408	-0.3429
Ground Truth	**tst**	**mwF1**	**mwF2**	**ht2**	**mwF2**	**ht2**	**mwF1**	**mwF2**
Hard Score	0	1	1	1	0	0	1	1
Soft Score	0	1	1	1	0	1	1	1

Appliance	**+90s**	**-95s**	**+100s**	**-105s**	**-110s**	**+120s**	**-140s**	**+160s**
fnH	0.2235	0.9845	0.2916	0.9791	0.8760	-0.2767	0.9882	-0.6207
fnM	0.3140	0.9760	-0.1120	0.9910	0.7598	0.0612	0.9986	-0.9299
fnL	0.2870	0.9744	-0.1516	0.9901	0.7586	0.0938	0.9971	-0.8434
ht3	0.2576	0.8941	-0.9652	0.8514	0.9921	0.8371	0.8621	0.0286
ht2	0.2594	0.9025	-0.9649	0.8614	0.9894	0.8365	0.8720	0.0257
ht1	0.2637	0.9807	-0.9607	0.9920	0.7867	0.8334	0.9999	0.0273
mwF1	-0.2135	0.0000	-0.0881	0.0000	0.0000	0.1048	0.0000	0.9889
mwF2	-0.0992	0.4282	-0.3015	0.5282	0.4838	0.3131	0.4992	-0.2216
mwI	0.0804	0.9734	0.8600	0.9572	0.9222	-0.7996	0.9663	-0.3806
frg	0.3622	0.9739	0.8802	0.9898	0.7591	-0.9699	0.9978	-0.0742
ktl	0.2800	0.8043	-0.9573	0.7453	0.9983	0.8263	0.7564	-0.0250
lmp	-0.0840	0.9763	0.9873	0.9621	0.9137	-0.9142	0.9717	-0.0525
snw	0.2535	0.9816	-0.9662	0.9919	0.7914	0.8389	0.9998	0.0370
tst	0.3845	0.9818	-0.1690	0.9916	0.7939	0.0980	0.9999	-0.8450
Ground Truth	**lmp**	**lmp**	**lmp**	**lmp**	**mwF2**	**ht3**	**tst**	**mwF1**
Hard Score	0	0	1	0	0	0	0	1
Soft Score	0	0	1	0	0	0	0	1

Appliance	**+162s**	**-167s**	**-170s**	**+180s**	**+181s**	**+190s**	**-197s**	**+210s**
fnH	-0.5587	0.8556	0.2796	-0.7363	-0.5333	-0.1809	0.8517	0.8154
fnM	-0.2619	0.7327	0.2696	-0.9557	-0.2433	0.1743	0.7276	0.4680
fnL	-0.1474	0.7316	0.2687	-0.7620	-0.1369	0.1838	0.7264	0.2871
ht3	0.3765	0.9862	0.2723	0.0278	0.3784	0.9994	0.9849	-0.3494
ht2	0.3743	0.9827	0.2756	0.0257	0.3763	0.9995	0.9813	-0.3466
ht1	0.3922	0.7609	0.2707	0.0386	0.3937	0.9919	0.7560	-0.3540
mwF1	0.2693	0.0000	0.0000	0.9605	0.2545	0.0626	0.0000	-0.4439
mwF2	0.5719	0.4709	-0.5785	-0.1068	0.5686	0.2033	0.4707	-0.4564
mwI	-0.6358	0.9057	0.2853	-0.4464	-0.6210	-0.7926	0.9025	0.7671
frg	-0.5683	0.7320	0.2618	-0.0845	-0.5661	-0.7877	0.7269	0.4988
ktl	0.3426	0.9998	0.2555	-0.0270	0.3457	0.9978	0.9999	-0.3045
lmp	-0.4845	0.8964	0.2814	-0.0603	-0.4854	-0.9805	0.8930	0.4438
snw	0.3824	0.7659	0.2722	0.0363	0.3843	0.9993	0.7610	-0.3571
tst	-0.3478	0.7685	0.2707	-0.8676	-0.3280	0.2629	0.7637	0.5427
Ground Truth	**mwF2**	**mwF2**	**ht3**	**mwF1**	**mwF2**	**ht1**	**mwF2**	**mwI**
Hard Score	1	0	0	1	1	0	0	0
Soft Score	1	0	0	1	1	1	0	0

Appliance	+230s	+232s	+240s	-247s	-260s	-270s	-280s	Total (%)
fnH	-0.6591	-0.5402	0.2731	0.9938	0.9910	0.9733	0.9393	
fnM	-0.9529	-0.2477	-0.0753	0.9923	0.9724	0.9940	0.8501	
fnL	-0.8199	-0.1378	-0.1515	0.9911	0.9708	0.9943	0.8488	
ht3	0.0030	0.3704	-0.9703	0.8886	0.9171	0.8301	0.9996	
ht2	0.0004	0.3683	-0.9698	0.8974	0.9242	0.8409	0.9999	
ht1	0.0082	0.3869	-0.9632	0.9962	0.9798	0.9931	0.8717	
mwF1	0.9889	0.2556	-0.1464	0.0000	0.0000	0.0000	0.0000	
mwF2	-0.1814	0.5743	-0.3118	0.5375	0.5169	0.4861	0.5090	
mwI	-0.3852	-0.6186	0.8481	0.9780	0.9840	0.9463	0.9704	
frg	-0.0583	-0.5551	0.8673	0.9918	0.9718	0.9935	0.8494	
ktl	-0.0511	0.3374	-0.9629	0.7916	0.8338	0.7176	0.9780	
lmp	-0.0307	-0.4776	0.9850	0.9822	0.9865	0.9527	0.9650	
snw	0.0115	0.3763	-0.9708	0.9966	0.9808	0.9932	0.8754	
tst	-0.8628	-0.3342	-0.1687	0.9968	0.9815	0.9923	0.8774	
Ground Truth	**mwF1**	**mwF2**	**lmp**	**mwF2**	**lmp**	**mwI**	**ht1**	
Hard Score	1	1	1	0	0	0	0	**43**
Soft Score	1	1	1	0	0	0	1	**53**

Appendix G

DETAILED DISCUSSION OF RESULTS

This appendix contains detailed discussion and analysis of the results produced by TLM, CPM and EEC for each of the combined measurements. The observations presented in Section 5.6 were drawn from the work contained within this appendix and are summarised below:

• The appliance steady-state combination vectors utilised for TLM and CPM contain favourable regions where the probability of obtaining a correct match is higher than in other regions of the vector, reducing the reliability of the results for these two techniques.

• The variability present in the measured total power levels has a significantly adverse impact upon the accuracy of the steady-state power level identifications made by TLM and CPM.

• The degree of distinctiveness exhibited by each waveform event edge for an appliance operational state change has a positive impact on the accuracy of the identifications made by EEC.

• The correlation results returned by EEC may at times not be distinct

enough for a definite 'best match' to a particular appliance operational state change event edge to be made, adversely affecting the accuracy of the technique.

The results for the TLM, CPM and EEC approaches may be found in Appendices D to F respectively, and the accompanying ground truth data in appendix B.

G.1 Combined Measurement 1

G.1.1 TLM Performance: 3/8 - 38% Accuracy

The toaster activation event occurring at 10s is incorrectly identified by TLM, being mistaken for a combination of three other appliances. The combined sum of the expected real power consumption levels for those three appliances differs from the measured steady-state by only 1W, whereas the expected real power consumption of the toaster varies from the measured steady-state by more than 21W. Thus it may be said that the TLM method has made a reasonable assignment, given the expected and measured power consumption levels. However a 21W discrepancy is not at all unprecedented. The measured power consumption levels may vary for a number reasons, ranging from the behaviour of other devices in the system to noise, power supply fluctuations and other usage dependent factors, such as resistivity and temperature effects. This illustrates a major failing of the TLM method, in that one combination of appliances can easily look like a another, a fact which is exacerbated by even minor variabilities between expected and measured power levels. The issue is further aggravated as more appliances are added into the circuit being measured, as this increases the probability that appliance state combinations with similar power consumption levels will be found.

All of the subsequent identification errors in the TLM pass may be ascribed to the same problem, where incorrect appliance state combinations are found to be closer to the actual measured level than those expected from the ground truth data. Of note amongst the correct identifications is the toaster event beginning at 30s. Here the appliance is correctly matched, yet it is misidentified in three other instances of solo operation. This points once more to the effects of the variability of measured power levels in situations that would be expected to yield identical readings. Furthermore, where the toaster is incorrectly identified, a different combination of appliances is assigned to the steady-state in each case. The laboratory experiment contains only eight rel-

atively simple appliances, yet sufficient potential power consumption matches have been generated to adversely affect the accuracy of this disaggregation approach. Given that any electrical circuit found within a household could easily be populated with a larger number of appliances, this lessens the feasibility of successfully implementing TLM in real-world NILM systems.

G.1.2 CPM Performance: 4/8 - 50% Accuracy

The CPM method performs better than TLM for this combined measurement, improving it by one binary point via the correct identification of the toaster and fan-high combination beginning at 60s. Whilst the solo toaster steady-states are just as inaccurately matched as in the TLM pass, more consistency is present in the incorrect identifications, with the same erroneous combination being assigned at 50s and 71s. This provides an indication of the reduction in the number of similar potential power consumption matches available for each measured level, which may be attributed to the reactive power phase refinement of the real power appliance state combination vector. However, the error that remains only serves to emphasise the problem of similar levels discussed in Section G.1.1. Despite the reduced size of the appliance state combination vector, the number of potential matches still remains too high for steady-state techniques such as TLM and CPM to be effective.

A partially correct identification may be considered to be entirely incorrect in this context. For example, at 20s the combination of toaster and fan-high is incorrectly selected ahead of the toaster and fan-low. CPM has not truly recognised the presence of the toaster as a component in the steady-state, as it is merely matching numbers with no other information attached. Thus the presence of the same appliance in both the ground truth data and the CPM output may be considered to be relatively coincidental. Certainly, the two results are not completely disconnected, as the power consumption of the toaster puts it within the right range for potential inclusion in the

output. However, this connection is relatively insignificant given the number of alternative appliance state combinations with similar power consumption levels.

G.1.3 EEC Performance: 6/8 - 75% Accuracy (Hard), 8/8 - 100% Accuracy (Soft)

EEC provided the best performance of the three techniques for this combined measurement, taking the hard score as the parameter of most interest. The perfect soft score was attained through correct identification of the fan in each instance, with some incorrect identifications of the various operating speeds proving the difference between the two scores. Unlike the steady-state techniques, EEC can be credited for these partially correct assignments, as all three of the fan setting waveforms possess similar positive edges. Thus, a characteristic particular to the appliance in question has been used to make the identification and it cannot be considered to be random. The toaster and fan waveforms have distinctive positive edges, which could well have contributed to the relatively high accuracy attained for this combined measurement. It may be noted that the two hard score errors occur on negative edge events, and that the waveforms for both of the appliances involved feature approximately square and negative edges that lack significant transient features. Thus this disaggregation pass presents an example of the effect of waveform event edge distinctiveness upon the performance of EEC.

The EEC results also provide an opportunity to discuss how the use of memory, as presented in Section 6.3, could prove beneficial to the technique. At 50s a fan-medium state negative edge is found, which EEC mistakenly ascribed to the fan-low state. However, if it was known that the fan-low state had not been activated previously, then that incorrect state could have been eliminated as a possible choice. A similar error occurs at 71s. These two instances serve to deprive the technique of a perfect hard score for this pass, and thus the use of memory may be seen to have the potential to significantly

improve EEC in such scenarios where the accuracy is already relatively high.

G.2 Combined Measurement 2

G.2.1 TLM Performance: 2/17 - 12% Accuracy

As with the previous TLM pass, discussed in Section G.1.1, appliance state combinations within a similar range of real power consumption values are confused throughout this set of results. For example, at 19s the toaster and microwave-inactive states, with a combined value of 770W, are confused for the fan-medium, lamp and snackwich states, with a combined value of 777W. These sorts of incorrect identifications are a consequence of the variability of measured power levels and the large number of potential matches within a small range of power values. The 7W discrepancy found between the expected and measured power level is minimal, only making up a very small percentage of the total range of power values, as presented in Table G.3. Thus this TLM pass reinforces the observations drawn from the first combined measurement, emphasising how difficult it is to make correct steady-state real power matches under these conditions.

TLM scored 38% accuracy for Combined Measurement 1, yet it only managed achieved 12% accuracy in this instance, a significant difference in performance. This may be explained by variability in the measured power levels, bearing in mind that TLM utilises the same appliance state combination vector for all passes. Combined Measurement 1 must contain steady-states that are closer to the individual measurement power levels than found in Combined Measurement 2. However, it should also be noted that the former consists of only 8 entries, compared to 17 in the latter. Thus any random element of success will carry far more weight in the first combined measurement than the second, and this factor may well have influenced the disparity between accuracies seen for the two TLM passes. This unequal weighting

of the individual disaggregation instances within each combined measurement make it necessary to use the EWA evaluation approach, as discussed in Section 5.5, if a clear picture of the overall technique accuracies is to be gained.

G.2.2 CPM Performance: 8/17 - 47% Accuracy

The microwave, which features heavily in this combined measurement, has a reactive power component that was only matched with 71% accuracy. All of the reactive power errors occur where the microwave-inactive state is operational at the same time as an appliance with medium to high real power consumption levels, such as the toaster, resulting in a large reduction of the measured reactive power level that leads to the steady-state being identified as a null state. This phenomenon, discussed in Sections 3.3.2, 3.3.3, 3.3.6 and 3.3.9, led to the investigation of an alternative CPM variant that excluded the microwave-inactive state from the reactive power matching phase. Section 5.3.1 presents the findings that resulted in its retention. At each point in time where this may be noted, such as at 19s, the microwave-inactive state is misidentified as a null state and its real power equivalent is removed from the appliance state combination vector. This prevents it from being rightfully identified during the real power matching phase, and thus enforces a incorrect identification for that steady-state period.

Despite this handicap, CPM still improves markedly upon the TLM score for this combined measurement, increasing the number of correct identifications by 6. Perfect scores are gained for every microwave state operating without the presence of other appliances. In these cases, the correct microwave state must be picked from a reduced real power appliance state combination vector that will only contain combinations that include that particular microwave state. Since the closest approximation to a null state in that vector will be the microwave state in question, the probability of a correct identification is vastly improved. These cases make up all of the additional 6 scores between

the TLM and CPM passes. Thus CPM has failed to improve identification accuracy when it comes to the more complex steady-states, providing a further illustration of how the reduced appliance state combination vectors are still too large for consistent accurate matching of levels to be realised.

G.2.3 EEC Performance: 10/17 - 59% Accuracy (Hard), 10/17 - 59% Accuracy (Soft)

The distinctive microwave positive edges are identified with 88% accuracy in this combined measurement, with only a single edge being misidentified at 93s. The high level of accuracy exhibited for this appliance contributes to the overall performance of the technique, making up 64% of the correct identifications for this pass. This reinforces the observation that the distinctiveness of the waveform edge shape has a strong impact on the effectiveness of EEC, as would be expected for such a transient edge event technique and is discussed in Section 5.6.

However, the toaster also has a relatively distinctive positive edge compared to the other appliances included in the experiment. The toaster waveform edge is not obscured by any other appliance activities, with well over 1000 data points to be found between it and the nearest magnatron event. This could provide a counter to the argument made above. However, it does not hold much weight as the toaster only occurs once in this measurement, and has a less distinctive shape than the microwave-front edges do. The toaster is mistaken for the kettle where it becomes active at 19s. Whilst the two appliances have very different peak amplitudes, they have moderately similar shapes, both possessing a transient peak and gradual ramp down to steady-state. This may have contributed to the kettle showing a higher correlation value than the toaster for this event.

Microwave-inactive events are incorrectly attributed to both fan-low and fan-high throughout the measurement. All three waveforms have relatively sim-

ilar peak amplitudes and edge shapes. At 17s the microwave magnatron negative edge is mistaken for the snackwich. Again, both have relatively similar negative edge shapes, but the magnitude of the change in each case is very different. All of the appliance waveform negative edges are relatively square and featureless compared to the level of distinctiveness exhibited between the positive edges, as shown in Chapter 3.2. This should theoretically lead to the accuracy of positive event edge identifications outperforming the negative edge matches for all of the combined measurements. The values contained in Table G.1 serve to confirm this, showing how positive edge EEC identifications outperform negative edges throughout the vast majority of the measurements.

Table G.1: Positive and negative event edge EEC accuracies for all combined measurements, showing hard and soft scoring.

	Hard Positive (%)	**Soft Positive (%)**	**Hard Negative (%)**	**Soft Negative (%)**
Comb. Meas. 1	100	100	50	100
Comb. Meas. 2	70	70	43	43
Comb. Meas. 3	78	89	18	27
Comb. Meas. 4	86	86	43	71
Comb. Meas. 5	50	75	25	38
Comb. Meas. 6	67	72	38	46
Comb. Meas. 7	14	29	0	0
Comb. Meas. 8	17	17	40	40
Comb. Meas. 9	72	78	0	15
Total	61	68	29	42

G.3 Combined Measurement 3

G.3.1 TLM Performance: 5/29 - 17% Accuracy

This combined measurement contains a null state midway through the ground truth data that is correctly identified, as are all null states found throughout the measurements. This provides an 'easy' score for TLM, as the closest appliance state level to the null found within the appliance state combination vector is fan-low, at over 25W. Thus noise exceeding this level would have to be present in the measurement in order for an incorrect identification to be made. The toaster state operating alone is correctly identified at 50s, yet TLM fails to match it accurately at 89 second. This provides another example of the inconsistency of TLM disaggregation performances throughout the combined measurement.

G.3.2 CPM Performance: 11/29 - 38% Accuracy

CPM improves significantly upon the TLM performance for this combined measurement, making an additional 6 correct identifications. The reactive power measurement contains multiple microwave events, which are matched with 83% accuracy, failing only in the instances where the microwave-inactive state is affected by the operations of other appliance. As discussed in Section G.2.2, CPM exhibits high accuracy when an appliance with a reactive power component is operating alone, given that this state becomes equivalent to a null state within the real power appliance combination vector that is produced, and is thus easy to match accurately.

An exception to this may be found at 152s, where the microwave-back state was operating alone, but CPM identified the fan-low state as being active as well. This indicates the presence of sufficient noise to raise the measured real power level to a point that better approximated the fan-low and microwave-back combination than the microwave-back state alone.

G.3.3 EEC Performance: 16/29 - 55% Accuracy (Hard), 19/29 - 66% Accuracy (Soft)

As with Combined Measurement 2, the presence of the readily matchable microwave-front edges in the combined measurement results in a better EEC performance than exhibited by the two steady-state techniques. In this pass the magnatron positive event edges were identified with 100% accuracy. The first positive toaster event edge is misidentified at 19s, but the second activation of the appliance at 169s is correctly identified. This reiterates the inconsistency of identifications for the same appliance within a single combined measurement, where the shape of the event edge of the appliance in question is not particularly distinctive.

G.4 Combined Measurement 4

G.4.1 TLM Performance: 8/14 - 57% Accuracy

The commonly observed types of power level matching errors may be seen in this set of results. For example, at 9s the snackwich is mistaken for a combination of the fridge, lamp and heater-low, where a steady-state power level of 698W is measured. The former has an individual steady-state level of 689W, and the latter combination consumes an expected power level of 700W. The last six identifications prior to the final null state are made correctly, all of which involve low real power consuming appliances. The lower the measured steady-state level, the less likely it is that a large number of appliance state combinations will be found that have a similar power consumption value, resulting in lower potential error when performing matches. Earlier in the combined measurement, between 9s and 130s, the measured power consumption levels are higher and the accuracy is far lower, with only a single correct identification found at 69s. This observation leads to the discussion of favourable matching regions within the appliance state combi-

nation vectors that is conducted in Section 5.6.

G.4.2 CPM Performance: 10/14 - 71% Accuracy

As with Combined Measurement 1, no appliances with reactive power components are found in the aggregated measurement. CPM still reduces the size of the appliance state combination vector in such cases by eliminating any states that include the fridge or microwave. As with the TLM pass, the snackwich and lamp combination found at 69s is correctly identified. But unlike the TLM pass, it is accurately matched again at 110s. A further improvement is found at 9s where the snackwich state operating alone is correctly identified. However, the larger combinations of appliances, which draw higher real power levels, remain misidentified. Thus whilst the reduction in the number of possible appliance state combinations has led to increased accuracy, the same underlying challenges facing TLM are still seen in the results for this CPM pass.

G.4.3 EEC Performance: 9/14 - 64% Accuracy (Hard), 11/14 - 79% Accuracy (Soft)

For this combined measurement EEC performs worse than CPM, although still offering better accuracy than TLM. This is the first of the combined measurements where it has not been the best of the disaggregation techniques. Of note in the results is the correct identification of both the snackwich positive and negative event edges. This appliance has a characteristic waveform where no steady-state is reached, with a steady ramp falling from the front to the back edges as shown in Section 3.2.5. This means that both edges are relatively distinctive, which is unusual amongst the appliances included in the experiment, most of which have waveforms with square and featureless negative edges. Thus this characteristic may have assisted EEC in correctly identifying both snackwich events, as seen at 9s and 144s.

The lamp is activated twice within the combined measurement, first at 29s and then again at 210s, with deactivation events following in each case. Only one of the two positive edges is correctly identified and none of the negative edges are matched to the right appliance. The positive edges of the lamp waveform are not distinctive, although there is a small ramp found directly after the initial edge, and the negative edges are square with no accompanying features. This influences the performance of EEC for this appliance, where only 25% of the edge events were accurately identified in this pass.

G.5 Combined Measurement 5

G.5.1 TLM Performance: 16/16 - 100% Accuracy

TLM turns in a perfect performance for this combined measurement, with every match up being made correctly. Given the poor performance of the technique for all of the previous combined measurements, this result is unexpected. However, a look at the power consumption levels of the appliances featured in the combined measurement provides insight into the mechanics behind the perfect accuracy. The heater-low state has the highest expected consumption, at 522W. Whilst this is not in the bottom range of appliance real power steady-state values, it is still only just above the mean for all the appliance states and ranked the 8th lowest power consumer, as shown in Table G.2. The lamp consumes 58W, and the fan states low through high consume only 26W, 30W and 38W respectively. This places the four states in the bottom 5 positions in terms of real power consumption.

The combination of appliances found in this combined measurement results in the measured power levels being low, which means that matches are being made towards the lower end of the appliance state combination vector. As discussed in Section 5.6, a favourable matching region in the appliance state combination vector may be found below the median levels. If a higher

Table G.2: Individual appliance steady-state real power consumption level statistics.

Appliance	**SS Power (W)**	**Ascending Rank**
Toaster	731.31	10
Kettle	2002.39	14
Lamp	58.05	5
Heater-Low	522.00	8
Heater-Medium	772.68	11
Heater-High	1291.06	13
Snackwich	689.39	9
Fan-Low	25.81	1
Fan-Medium	29.65	2
Fan-High	37.56	3
Microwave-Front	205.81	7
Microwave-Back	1063.30	12
Microwave-Inactive	38.73	4
Fridge	119.07	6
Mean	**541.91**	
Median	**363.91**	

heater level were used it could result in a lower accuracy being realised, as a larger number of appliance combinations would be available to be mistaken for the measured power level. Examples of this may be seen in combined measurements 6, 7 and 8. Table G.3 shows the maximum measured power level for each of the combined measurements, along with the highest correctly matched power level for TLM and CPM.

Combined measurements 4 and 5 have the lowest measured steady-state levels, and the best accuracies for both TLM and CPM. This relationship does not hold for all of the measurements, but it is interesting to note the significantly lower measured steady-state power levels and the corresponding markedly superior accuracies in these two cases. The total expected power consumption of the heater-low, fan-high and lamp states is 617W, which is well below the closest expected individual appliance power level of 689W for

Table G.3: Maximum measured and correctly matched steady-state real power levels for TLM and CPM.

	Maximum Steady-State Measured (W)	**Highest Correct TLM Match (W)**	**TLM Accuracy (%)**	**Highest Correct CPM Match (W)**	**CPM Accuracy (%)**
Comb. Meas. 1	766.65	760.95	38	768.87	50
Comb. Meas. 2	1748.70	38.73	12	1063.30	47
Comb. Meas. 3	1842.69	770.03	17	1063.30	38
Comb. Meas. 4	778.23	747.44	57	747.44	71
Comb. Meas. 5	617.06	617.61	100	617.61	100
Comb. Meas. 6	2547.29	324.88	16	1063.30	32
Comb. Meas. 7	2085.08	789.35	43	1349.10	64
Comb. Meas. 8	2212.47	2002.39	9	2002.39	9
Comb. Meas. 9	2567.62	937.12	23	1585.30	32

the snackwich. The fridge and heater-low states give a combined level of 641W, which is still removed from 617W. The closest appliance state combination vector power level found below 617W is 459W, generated by the fridge, microwave-front, fan-high and lamp states. Looking through the TLM results for all of the combined measurements, it may be seen that none of these numbers are likely to be matched to one another.

This provides another illustration of the manner in which certain appliance combination steady-state levels can sit in a region of the appliance state combination vector where the probability of a correct match being made is radically increased, due to the lack of similar levels in the vector. Section 5.6 contains further discussion of the existence of favourable matching regions, the range of values of which are dependent on the power characteristics of the appliances incorporated in the circuit being disaggregated.

G.5.2 CPM Performance: 16/16 - 100% Accuracy

As the CPM method may be seen to improve upon the TLM accuracy for each of the combined measurements, it may be assumed that CPM would score 100% for this measurement, being a refinement of the TLM method. However, this assumption is false. Where an appliance state is incorrectly matched in the reactive power phase of CPM, errors will be carried through to the real power matching phase. Thus could lead to CPM exhibiting poorer performances than TLM, although this is not the case for any of the combined measurements. However, no errors are experienced in the reactive power matching phase for this combined measurement, and thus CPM performs as expected. If the unrefined TLM real power appliance state vector generated by TLM performs at 100% accuracy, then the reduced version produced by CPM would be expected to match that performance under these conditions.

G.5.3 EEC Performance: 6/16 - 38% Accuracy (Hard), 9/16 - 56% Accuracy (Soft)

In this EEC pass, all of the fan-high, fan-medium and fan-low waveform positive edges are correctly identified. However, only 25% of the corresponding negative edges are matched accurately. The positive edges of the fan waveforms are relatively distinctive, possessing small transient peaks followed by ramps down to a steady-state, as presented in Section 3.2.6. As with the majority of the appliances, the waveform negative edges are square and conventional. Thus this performance of EEC is consistent with previous observations in terms of the disaggregation of fan events.

At 170s and 290s, both negative event edges, the true appliance state changes are mistakenly attributed to the snackwich. Given that the snackwich is the only appliance to feature a relatively distinctive negative edge, it is contrary to expectation that it should be selected ahead of other appliances with more conventionally shaped negative edges. A further unexpected error may be found at 30s, where an activation of the lamp is attributed to the fridge. The fridge has a very distinctive waveform positive edge, as presented in Section 3.2.8, that markedly it from that of the lamp, making it a particularly counter-intuitive match.

None of the lamp and heater-low waveform edges are correctly identified where they are found in the combined measurement, regardless of their direction of change. As they make up 50% of the detected events, this particular poor performance makes a significant contribution towards bringing down the overall accuracy for this combined measurement. Both the lamp and heater have unremarkable positive and negative waveform edges, which may contribute to their repeatedly incorrect identification here.

G.6 Combined Measurement 6

G.6.1 TLM Performance: 5/31 - 16% Accuracy

This combined measurement includes the states heater-medium and heater-high, which have expected power consumption levels of 773W and 1291W respectively. None of the steady-states containing either of these two appliance states are identified correctly. The higher power consumption of these two states most likely places the combinations that they are included in into a densely populated range of the appliance state combination vector, making it more difficult to obtain an accurate match. For example, at 80s the ground truth data shows the following appliance combination; fridge, heater-medium and microwave-front with a combined expected power consumption of 1098W. However, the vector entry matched to this steady-state contains the fan-low, microwave-front, fridge, lamp and snackwich states with the same 1098W power consumption level, where the measured steady-state is 1099W. This provides a classic example of the fundamental flaw found in TLM, that even a small number of appliances can generate similar or even approximately identical steady-state power levels, making accurate matching implausible under non-ideal conditions.

The few correct identifications made in this TLM pass feature combinations of relatively low power appliances. Examples of this may be seen at 193s and 278s, where the measured steady-states levels are 167W and 315W respectively. These successful identifications complement the favourable matching region discussion conducted in Section 5.6. However, numerous other instances of low measured power steady-states are incorrectly identified throughout the measurement, such as the 38W level at 199s and the 168W level at 259s. Thus, whilst the concept of favourable matching regions still holds, a low measured steady-state level is no guarantee of an accurate match being made.

G.6.2 CPM Performance: 10/31 - 32% Accuracy

As might be expected, CPM improves upon the TLM results for this disaggregation pass. The reactive power matching phase is performed with 90% accuracy, where the 10% shortfall may be predominately attributed to the incorrect identification of microwave-inactive states in the presence of large real power consumption. A rare reactive power matching error not involving the microwave-inactive state may also be seen at 181s , where the combination of fridge and microwave-back states is taken to be the microwave-back alone. This could indicate the presence of high reactive power noise for this steady-state period, or could be an occurrence of the same reactive power measurement inconsistency discussed in Section 3.3.4.

Similarly to the CPM performance for Combined Measurement 3, the main source of the improvement for this combined measurement comes from the correct identification of the appliances with reactive power components when operating without the presence of other purely resistive appliances. Thus the reactive power matching phase serves to introduce easily matchable artificial null states into the results, as discussed in Section G.3.2. The remaining error for this combined measurement may be ascribed to the same steady-state level identification difficulties discussed for TLM and CPM earlier in this appendix.

G.6.3 EEC Performance: 17/31 - 55% Accuracy (Hard), 61% Accuracy (Soft)

This combined measurement contains microwave events, which are well identified by EEC. For this pass, 79% of the microwave positive waveform edges are accurately assigned. The corresponding negative edges are only identified with 50% accuracy, as fits previous EEC passes containing this appliance. The more distinctively shaped microwave positive edges are easier for EEC to match than the square and conventional negative edges, complementing

the discussion conducted in Section 5.6. This observation may be applied to the relatively square heater edges, which lack distinctive transient features, and are only identified with 33% accuracy in this combined measurement.

However, the fridge is not correctly identified at any point in the combined measurement. Given the distinctive shape of the positive waveform edge for this appliance, presented in Section 3.2.8, and the accuracy exhibited by EEC when matching the similarly distinctive microwave positive edges, a relatively high level of accuracy would be expected here. Throughout the combined measurements, six fridge events may be found, which are split equally between positive and negative edges. None of these events are correctly identified. This poor performance contradicts the fundamental underpinnings of EEC, as the distinctively shaped edge event of the appliance has not led to good identification performance.

This poor accuracy may be ascribed to two factors. Firstly, the large transient peak found on the positive edge is relatively short in duration and thus may have minimal impact on the overall shape of extracted edge, depending on the sample parameters utilised. Where EEC was applied using a positive edge sample length of 75 data points, 30% of the length used here, a single identification was performed correctly. However, the match was made for a negative edge, and thus does not necessarily support the argument that the shortening of the sample length emphasises the distinctiveness of the edge event shape. Secondly, the fridge waveform negative edges are square and conventional in shape. Thus the three positive edges found throughout the combined measurements, with their pronounced transient peaks, are of the most interest in the context of this discussion. This means that conclusions drawn from the incorrect identifications of the fridge do not carry sufficient weight to be extended to EEC as a whole, given that such a small sample size could easily be subject to other random factors that might render its performance anomalous.

G.7 Combined Measurement 7

G.7.1 TLM Performance: 6/14 - 43% Accuracy

Two of the correct identifications made during this TLM pass are for the lamp operating alone, at 10s and 229s, where this appliance state combination is matched with 100% accuracy throughout the combined measurement. Whilst this appliance has a low power consumption, 58W, it is surrounded by a few close potential matches due to the fan consumption levels. For example, a combination of the fan-low and fan-medium states consumes 55W, and fan-high with fan-low consumes 63W. Thus the accurate matching of these solo lamp states indicates the presence of low noise levels in this combined measurement.

A null state may be found at 135s. The two null and lamp steady-states featured in the combined measurement make up 4 out of 6 of the correct identifications, a significant portion of the TLM score for this pass. TLM identifies null states relatively easily, given that noise of over 13W would have to be experienced in order for it to be mistaken for the fan-low level of 26W. Whilst such a noise level would not be unprecedented, it is unlikely to be found during a null state, as appliance operations are responsible for much of the variation found between measured and expected power values.

G.7.2 CPM Performance: 9/14 - 64% Accuracy

This combined measurement includes no appliances with reactive power components, thus CPM limits the appliance state combination real power vector for all steady-states in the combined measurement to combinations of purely resistive appliances. The TLM pass for this combined measurement included many appliances with reactive power components in the identifications, each of which were logically erroneous. Thus this limitation of the state combination vector must have positive implications for the disaggregation perfor-

mance of CPM, as may be seen in the 21% increase in accuracy between the TLM and CPM passes. Both the solo toaster and heater-high states, at 129s and 149s respectively, are misidentified. This reinforces the idea that appliance steady-states featuring higher power consumption levels are more difficult to match accurately, as they do not fall into the favourable matching region discussed in section 5.6.

especially when considered alongside the TLM lamp identification performance discussed in Section G.7.1.

G.7.3 EEC Performance: 1/14 - 7% Accuracy (Hard), 2/14 - 14% Accuracy (Soft)

The poor performance exhibited by EEC for this combined measurement is sufficiently anomalous for it to be considered an outlier. Apart from the toaster, none of the event edges produced by the appliance states operated within this combined measurement are particularly distinctive. The heater and lamp have approximately square and conventional positive edges, which hinders the ability of EEC to distinguish between appliances accurately. As discussed earlier in this appendix, none of the appliances have distinctive negative edges except the snackwich, which is not included here. Thus the poor performance of EEC for this pass fits in with the discussion of event edge distinctiveness conducted in Section 5.6, given that the single correct identification is for a toaster positive edge, found at 193s, which is the most distinctive of the edges found in this combined measurement.

Another major problem hampering EEC performance throughout the combined measurements is the similarity between correlation results returned for many event edges. For example, at 90s the difference between the correlation figures for the correct and incorrectly matched appliance states is only 0.001. It is difficult to be confident about assignments made where the differences are this low. However, it should be borne in mind that the field of options

has already been significantly narrowed in these cases, so the final outcome cannot be considered to be entirely random. Conversely, EEC sometimes returns low correlation values for all appliance states at a particular edge. An example may be found at 169s in this combined measurement, where the highest correlation value was 0.43, the best of a number of bad matches. For both of these extremes, a correct match does not inspire confidence. A more convincing correlation value, differentiated from the rest of the field by a significant margin, would present a far more definitive outcome in such cases, as discussed further in Section 5.6.

G.8 Combined Measurement 8

G.8.1 TLM Performance: 1/11 - 9% Accuracy

This is the only combined measurement not to terminate in a null state. As null states are easily identified by the steady-state disaggregation techniques, this makes this combined measurement harder to aggregate for TLM and CPM. However, despite this impediment, the final steady-state is correctly identified as the kettle operating alone. This is the single accurate match made for this TLM pass. As the kettle is active for large portions of the combined measurement, the measured power levels are elevated for the majority of the steady-states. By virtue of being the highest power consuming appliance, at 2002W, all matches including the kettle must be made a relatively high power level within the state combination vector. This is associated with a low level of disaggregation accuracy, as discussed throughout this appendix, and in Section 5.6.

For example, at 49s the fridge, kettle and lamp are active with a combined expected power level of 2180W. However, the measured level for that steady-state is 2212W, and the resulting incorrect match is made at approximately the same power level. There are sufficient appliance state combinations avail-

able within that range for the measured level to be easily assigned to another erroneous appliance state combination vector entry, as has occurred here, where the 32W disparity between the measured and expected steady-state power levels is to large for a correct match to be a feasible outcome.

The single correct identification made for the solo kettle state occurs where the fridge is deactivated, as may be seen between 100s and the end of the measurement. Given the poor performance TLM exhibits throughout the combined measurements when matching the higher power consuming appliances, the kettle might be expected to be among the misidentified appliances in this pass. Taking this in mind, it is prudent to look at the fridge a little closer. At 79s, the total measured steady-state level is 150W, which is 31W higher than expected for the fridge when operating alone. This either means that the measurement contains high noise levels at this point, or that the fridge is simply consuming unexpected levels of power. It is possible that the fridge entered some unanticipated mode of operation that was not present during the initial measurements, resulting in the additional power consumption. This raises a further concern involving the use of steady-state power disaggregation techniques such as TLM and CPM; appliances may possess hidden states that become active at unanticipated points in the total power measurements, introducing error into the disaggregation process.

G.8.2 CPM Performance: 1/11 - 9% Accuracy

In a rare state of affairs, CPM fails to better the reported accuracy of TLM for this pass. The fridge, which is almost constantly present throughout the combined measurement, possesses a reactive power component which is identified with 100% accuracy. This means that the appliance state combination vector is reduced without introducing any error into the results, and thus this CPM pass might well be expected to improve upon the TLM performance significantly.

Whilst the scores are not improved, perhaps as a result of the high power

levels and possible variability of fridge operational levels discussed for TLM in Section G.8.1, the erroneously assigned state combinations do differ between the TLM and CPM results. At 9s, 39s, 49s and 64s the combinations matched by CPM to the measured levels are closer than seen in the TLM pass. Thus CPM can be seen to have provided some form of improvement in this case, however marginal.

G.8.3 EEC Performance: 3/11 - 27% Accuracy (Hard), 3/11 - 27% Accuracy (Soft)

Whilst EEC does outperform TLM and CPM here, it still delivers one of its worst accuracies. The three successes arise from the correct identification of the positive and negative edges of the kettle waveform, plus the negative edge of the lamp waveform. None of these event edges are particularly distinctive, and thus they would not necessarily be expected to be candidates for accurate matching. The fridge is the only appliance featured in this combined measurement that might be expected to be readily correctly identified by EEC, given its distinctive waveform positive edge. However, as discussed in Section G.6.3, the positive and negative waveform edges of this appliance are matched poorly throughout the combined measurements. Aside from the notable exception of the fridge, the low accuracy realised when attempting to apply EEC to the set of appliances included in this combined measurement are a product of their lack of distinctiveness, as discussed earlier in this appendix and in Section 5.6.

G.9 Combined Measurement 9

G.9.1 TLM Performance: 7/31 - 23% Accuracy

Whilst this combined measurement contains the highest number of appliances, there is no point in time where all four appliances are simultaneously operational, as discussed in Section 3.3.9. A number of high power consumption states are included amongst those present in the measurement, namely the heater-high, microwave-back, heater-medium and toaster states, with expected consumption levels of 1291W, 1063W, 773W and 731W respectively. These levels serve to push the measured real power steady-state values into the mid to upper range on the appliance state combination vector, making it hard for TLM to make accurate matches. Examination of the differences between the measured and matched real power levels reveals that the majority of matches are close in value. This indicates that the large number of appliance state combination levels available for matching in the higher ranges of the state combination vector are largely responsible for the low level of accuracy realised by TLM for this pass. This is in line with the related discussions conducted throughout this appendix and in Section 5.6.

G.9.2 CPM Performance: 10/31 - 32% Accuracy

As might be expected, CPM improves on TLM for this combined measurement by reducing the size of the appliance state combination vector. The reactive power matching phase is only conducted with 87% accuracy, thus some error is introduced into the process that prevents CPM from providing a greater improvement on TLM than the 9% reported here. The main source of the error found amongst the reactive power identifications is the erroneous identification of the microwave-inactive state as a null state, which is a side-effect of the CPM variant chosen, see Section 5.3.1. However, a variation on this commonly observed CPM error is seen at 210s, where the

microwave-inactive state is identified as being present in the reactive power measurement even though the ground truth data indicates that this is not the case. The measured reactive power steady-state at this point in time is 18VAR, which is the cause of this error, and could be ascribed to noise or some other variation in the reactive power measurement.

G.9.3 EEC Performance: 13/31 - 42% Accuracy (Hard), 16/31 - 52% Accuracy (Soft)

Whilst this is not a great performance for EEC, it nonetheless provides a marked improvement on the accuracies of TLM and CPM. A big factor in this success is the presence of the microwave in the combined measurement, with 100% of the distinctive microwave positive edges being identified correctly. None of the microwave negative edges, which are approximately square and without significant transient features, are accurately matched. This reinforces the observations drawn from the discussions of event edge distinctiveness conducted for EEC throughout this appendix and presented in Section 5.6. For the rest of the event edges in the combined measurement, the lack of both distinctive transient features and definitive correlation results serve to keep the each accuracy low.

With the microwave positive edge events removed from the combined measurement, the overall accuracy falls to 5/23, or 22%. This highlights a positive attribute of EEC, as the performance is indicative of how reliant the technique is upon the individual characteristics of the appliances included in the total power measurement. Thus, any improvement introduced into the detection of those characteristics, such as the extraction and emphasis of the most distinctive shapes from the sampled edges of the appliance waveforms, will increase the accuracy of EEC for all of the appliance states that it encounters.

Appendix H

HCII2014 PUBLICATION

A paper covering early research into the development of an audio warning system for power consumption in rural South African households, which provided the initial inspiration for the research contained in this dissertation, was published in the HCII2014 conference proceedings and is presented in this appendix.

The residences targeted by this research feature pre-paid power meters, containing relatively low rated current breakers that frequently trip during peak power usage periods. The use of audio cues is investigated as a means of dynamically notifying residents of their electricity usage levels, both preventing failures and informing their power consumption behaviour.

A variety of audio characteristics, including tempo, rhythm, pitch and volume, are assessed; such that the best metrics may be established for designing audio cues to be implemented in this context.

User-Centred Design of an Audio Feedback System for Power Demand Management

Rebecca Ford[1], Joe Penn[2], Yu-Chieh Liu[2],
Ken Nixon[2], Willie Cronje[2], Malcolm McCulloch[3]

[1] School of Engineering and Computer Science Victoria University of Wellington
PO Box 600, Wellington, 6140, New Zealand
rebecca.ford@ecs.vuw.ac.nz
[2] School of Electrical and Information Engineering,
University of the Witwatersrand, Johannesburg, Private Bag 3, WITS, 2050, South Africa
joseph.penn@students.wits.ac.za,
{yu-chieh.liu,ken.nixon,willie.cronje}@wits.ac.za
[3] Department of Engineering Science, University of Oxford
Parks Road, Oxford, OX1 3PJ, United Kingdom
malcolm.mcculloch@eng.ox.ac.uk

Abstract. Low cost houses in South Africa are supplied with a pre-payment meter and a circuit breaker that trips at a low power level (about 20A, 4.5kW), resulting in many nuisance trips. Four categories of audio cues, each being able to represent five levels of power consumption, are assessed. A survey of 62 people was conducted. The numerical analysis of the results and the perceptions of the respondents both indicate that the use of changing tempo and texture is the most effective at conveying feedback information on the power consumption in the home.

Keywords: audio cues, power demand feedback, low cost

1 Introduction

This paper addresses the issue of developing a design methodology for providing immediate and intuitive audio feedback about high power consumption, especially for periods when the power level is approaching the capacity of the main circuit breaker.

On any electrical power system (national grid, microgrid or nanogrid) it is extremely important that the flow of power between generators and loads is balanced at any instant in time. This ensures stable operation of the system and avoids the disruption that will ensue if the grid is blacked out due to instability.

Stability can be addressed from the generation side as well as the consumption side. An adequate reserve margin on the generation side (embodied in the kinetic energy of the spinning turbo-generators, or stored battery charge on microgrids) gives the grid operators the freedom to dispatch more energy from the generators to the load side at short notice. In particular, South Africa is facing severe generation constraints at the present moment in time. The generation reserve margin of the national utility company (ESKOM) has been as low as 0.17% on 13 May 2013 [1]!

Load side response (better known as demand side management) is now coming to the fore, as it has been demonstrated that it can be more economical than expanding the generation side [2]. However, demand side response is challenging because it requires that a large number of consumers actively participate. In South Africa, national campaigns are in place to encourage households to swap incandescent lights for more efficient lighting solutions and consumers are being offered rebates on solar water heaters [3]. Furthermore, real time alerts are displayed on state-owned television channels to reduce peak demand; this visual information system takes the form of a special graphical display at the bottom of the television screen that indicates the current demand status to households via the use of colours and bar charts. The scheme provides information to consumers about the state of the grid, and has been shown to have an impact at a national level [4], but does not tell consumers much about their own contribution to the total demand. This is a problem because many residents in rural areas often have their power consumption limited by pre-paid electricity meters, which are equipped with feed-in breakers that trip at a modest level of 20A [5], cutting off the power supply with no warning.

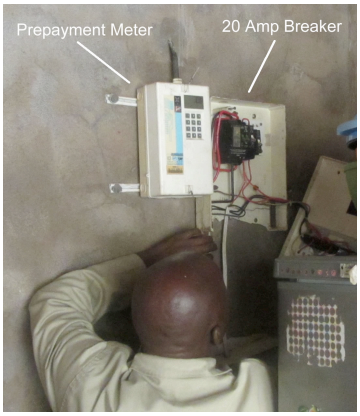

Fig. 1. A typical installation showing the pre-payment meter and 20A breaker.

Although there is the occasional use of automation to disconnect hot water systems during periods of high power demand [6], this intervention is often not sufficient to prevent the breaker from tripping. Further intervention is frequently necessary, but the automation of additional household appliances becomes complex and is too expensive, especially given that most residences equipped with prepayment meters and feed-in breakers are low-income households.

However, site visits revealed that the combination of their low income and use of prepayment meters has made these residents both aware of their household energy usage and motivated to take action to reduce consumption and to prevent tripping of the breaker. As there is a strong intrinsic motivation amongst the community [7] this context presents an ideal case for integrating users into the demand management process. The feedback about household electricity demand is thus explored as a mecha-

nism for encouraging and enabling users to better manage their energy consumption and prevent power outages.

This paper focuses on assessing the efficacy of cues on the user to enable them to manage their demand, via limiting high levels of energy consumption and prompting immediate action when the load approaches the trip level. Specifically, it addresses the question of which parameters within the audio cues produce consistent, accurate and meaningful responses from users. The work presented here does not include the deployment of any technologies into the field.

This paper explores the relevant literature to determine key aspects in the design of effective feedback mechanisms for demand management. Section 2 concludes that, in general, user-centric design needs to fulfill four criteria. The specific requirement of the users considered in this paper is that they need to respond immediately to prevent power outages, and thus the case for an audio cue is made in Section 3. However the users also want to be made aware of high power consumption, therefore a suite of cues is required. The choice of the design of the suite is discussed in Section 4. The efficacy of the design options are explored via the use of a survey in Section 5. The results in Section 6 highlight that two modalities are more effective than one.

2 Feedback as a Mechanism for Demand Management

Feedback about energy consumption has been used over the past 40 years as an effective mechanism for encouraging management of energy demand. Feedback interventions are on the whole effective at encouraging users to reduce overall consumption, and they are cost effective when compared to other interventions [8]. However, the way in which users respond to feedback about their consumption varies significantly, and whilst feedback is effective on average, it is not so in all cases [9]. A more recent body of work in this space points to the importance of considering users when designing feedback interventions, particularly with regard to their interaction with the feedback technology [10,11].

Although the provision of energy consumption information is of considerable value, for the feedback system to be effective at bringing about the desired shift in energy behaviour it is important that the design process accounts for the way in which users interpret and respond to the feedback, as well as their behavioural and motivational psychological aspects in relation to energy use [12]. In addition, contextual constraints can limit a person's ability to respond to feedback regardless of their motivation to act [13], and therefore careful consideration of the specific purpose of the feedback, the context in which energy is being consumed, and the Living Standards Measure grouping and cultural background of the target demographic is important.

A key challenge is to develop a user-centric design of a system capable of providing households with real-time feedback about their consumption that meets the following four criteria:

(1) is appropriate to the specific context in which it is intended to be used,
(2) is interpreted consistently and accurately,
(3) provokes a response at the appropriate point in time, and
(4) does not overburden or confuse the user.

The context for this study is different from most feedback studies and therefore needs closer examination and is addressed in the next section.

3 The Case for a Audio Cue

Most feedback interventions are designed to encourage consumers to reduce their overall energy consumption, and are designed and evaluated accordingly. However, these systems are not appropriate when trying to encourage reductions in peak usage, where the main concern lies around the simultaneous use of three or more high power appliances leading to a power trip. There are two significant implications.

Firstly the user response has to be immediate. Energy consumption is a measure of power demand aggregated over time and hence the timing of the feedback to the user is not critical. For this case the user can 'pull' the feedback from the device. However for peak power response the feedback must reach the user immediately, hence the device must 'push' the information to the user.

Secondly *all* users in the home must be aware of the feedback, no matter where they are located within the home, as they may each cause the breaker to trip by increasing the load. Hence the feedback must not be a point source of information but rather have a ubiquitous reach.

A third issue is that the feedback needs to indicate the level power used, ranging from moderate to extreme. The reason for this is that, even at the moderate level (e.g. just the oven on), the addition of just a single further high power device (e.g. iron) and one medium power device (e.g. fridge) can lead to a trip.

A fourth context related issue is that the target community is low paid, and therefore cost is a constraint on the implementation.

Typically feedback is provided to users visually. A user information [13] unit has been trialled in South Africa where the user interface is a three colour (green, amber red) visual display. The feedback is triggered both by local measurements and from information communicated from a central control room. The drawback of the visual display is that the user is not always facing the information unit, or is perhaps not even in the same room. The product is intended for the utility who will own and operate it to manage load/demand.

Whilst visual displays have the potential to provide detailed information about electricity demand, they are not always located such that they are visible to the consumer at the necessary point in time. As users are often physically occupied with tasks that might increase their energy consumption, such as housework, they are unlikely to pay constant attention to the display. However, an audio cue can offer a superior alternative interface [14] that reaches a greater area of coverage in the house, and provides immediate notification of usage status to the consumer, thus addressing issues one and two. Issue three can be achieved through the use of a range of cues provided via interactive technologies, though care must be taken to ensure that they do not become a nuisance to users [15]. As audio devices are low cost, issue four is also solved.

Therefore the authors propose that the most appropriate form of feedback applicable to the specific context of power management (criterion 1) is that of a suite of au-

dio cues. The next section explores the idea that appropriately designed audio feedback can improve the level of positive responses from the end users. In addition this paper investigates whether different sound symbols can be used to effectively warn end-users about power as well as energy constraint. It also tries to determine if the audio symbols can be used to communicate a sense of the urgency of the problem.

4 Design Parameters of Use of Audio Cues

For the audio cues to evoke their intended responses in users, they must be readily distinguishable from one another across the various levels of energy usage to which alarm signals have been allocated. The implemented audio cue set must also be relatively intuitive to respond to, requiring a minimal learning period for users to become accustomed to the scale of intensity contained within the batch of samples. To achieve this the musical parameters are progressively increased corresponding to the increased power usage. However, the individual parameters for variation must be carefully identified in order to accommodate both the distinguishability and intuition requirements of the design so as to meet criterion 2.

When choosing audio parameters to investigate, it is crucial to consider the impact that they might have upon the user once introduced into their domestic environment. The audio samples utilised must induce sufficient annoyance at the critical end of the scale to bring about alterations to user behaviour, yet must also be benign enough to avoid excessive irritation for lower energy usage levels [16]. If the audio cues are too annoying at all energy usage levels, users will be inclined to eliminate the audio functionality of the energy monitors entirely. In order to achieve this aim, the sound samples must increase in 'urgency' or irritation factor by changing certain properties as the level of energy usage increases.

Although certain elements (such as melody, harmony and rhythm) may be used to impart levels of urgency, they offer consistent irritation levels to users and thus are not appropriate for this application. For example if a major-harmony themed melody is repeated over a sustained period it may impart less urgency than a minor-harmony themed alternative [16]. However, the constantly looping phrase is likely to be equally irritating to the user regardless of the variant, quite possibly resulting in deactivation of the monitoring device entirely. Accordingly, foundational musical elements that can be utilised with simple tones in order to create audio cues that feature high degrees of fundamental variation offer the best building blocks for the sonic elements required for this application.

The pitch (or 'frequency') of a note is a fundamental musical property that can be varied with profound effect. Some people struggle to recognise subtle fluctuations in tones all are capable of recognising substantial changes in pitch. As lower pitched tones sit quite subtly amongst background noise and higher pitched tones tend to cut through more noticeably, variations in pitch are an ideal parameter to explore in this domestic context.

Tempo (or 'speed') is one of the most basic musical devices, variations of which are instantly recognisable. As tempo is entirely independent of pitch, it may be recognised and experienced by even the most 'tone-deaf' and musically-uneducated

amongst us. Furthermore, different tempos are distinguishable from one another, making tempo a natural element to be exploited in this application. Extremely slow tempos can result in long intervals between sonic elements, reducing the irritation and urgency factor associated with an audio cue. Fast tempos have the opposite effect, and hence its efficacy is further explored.

Texture is the tactile quality that may be ascribed to a sonic element, an abstract concept that often leads to the use of adjectives such as 'rough', 'smooth', 'round' or 'thin' in order to describe sounds. It is a fundamental building block of music, and is easily distinguishable to the human ear, being entirely separate from harmony. Given that the textures of sounds can have effects on listeners that range from 'soothing' to 'jarring', this element is a natural candidate for inclusion in the application in question. However as the variations are more subtle, it was combined with changing tempo.

In addition to such musically-oriented parameters, we are subjected to a diverse range of audio stimulus that effects our behaviour, such as the hooting of car horns, barking of dogs and so forth. Accordingly, the use of such audio can be used in order to generate responses in people that are directly related to generic experiences of the world around us and do not require any level of musical abilities in order to distinguish. This makes the use of non-musical sonic samples, recorded from the surrounding environment worthy of investigation.

Thus the properties of pitch, tempo, texture and real-world association were chosen for evaluation in this application. The sonic samples utilised for the real-world association category of audio cues were selected from within an animal theme, using fairly generic animal sources. The noises selected for use were deemed to be both fairly universal (mainly domestic animals) and to provide a subset of sounds to which the vast majority of users would have been exposed with relatively high frequency during their lifetimes. To see how users would react to the sounds a survey was conducted.

5 Survey Methodology

The purpose of the survey was to determine how effective each of the four categories can be distinguished to represent the following five levels of power consumption:

(a) Moderate power usage: above average rate of consumption.

(b) Moderate-high power usage: significantly above average.

(c) High usage: energy consumption should not be increased further

(d) Very high power usage: approaching trip level of main breaker, reduce usage as soon as possible.

(e) Extreme power usage: about to trip main breaker, immediate action required.

A ten second sound sample was generated for each level. The musical properties of interest were incrementally increased for each sound representing the correspondingly increasing power level.

A set of 15 randomly selected sound samples per category was placed in a video. The first five samples randomly covered all 5 levels. We call this the 'learning stage' as this is the first time the person is exposed to the sounds. The next 10 samples randomly covered each of the 5 levels twice. This latter data is evaluated for consistency

and accuracy. As each sample was played, the person was asked to identify which level they thought the sound represented.

To avoid bias users were not told how the sounds may vary and the categories were randomly presented to the users.

At the end of the survey, users were then asked two open questions: (1) which category they thought the most effective and (2) at which point they would take action. Anonymous demographic information was also collected.

6 Results

There were 61 respondents (8 New Zealand, 21 Southern Africa, 32 UK; 19 female and 34 under the age of twenty and 10 over the age of 50). The mean time to complete the survey was 17 minutes. Three respondents that did not complete the survey were discarded.

6.1 Criterion Two

The second criterion is that the feedback cue is interpreted consistently and accurately. To test this using the survey data, consistency is measured using the metric of the percentage of users whose second and third responses to the same sound level were identical, Fig. 2. Accuracy represented by the offset between the actual and the perceived level. The metrics used are the mean and standard deviations of this offset, Fig. 3.

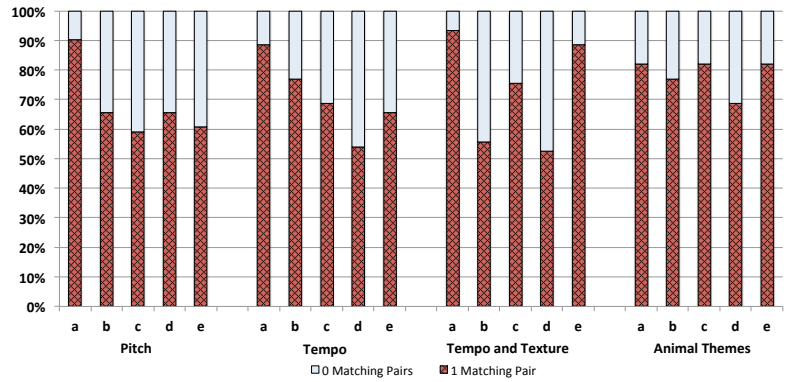

Fig. 2. Consistency of user responses to second and third iterations of each audio cue level

Pitch

Consistency: Less than 66% of the respondents are consistent in 4 of the 5 levels. The exception is level (a) which has 90%. However, it should be noted that the low frequency sample used for this set of audio cues could not be heard clearly on many

laptop and cellphone audio-speakers, thus appearing as a change in volume and pitch, which can explain the high level of consistency. This category had the worst overall results for consistency.

Accuracy: The standard deviation and offset plots show a distinct worsening of performance as the pitch is increased. At level (e) the offset is >-1.5 indicating that on average all the respondents severely underestimated the urgency. This category again had the worst overall results for accuracy.

Tempo

Consistency: This category featured high consistency levels for the lower levels, with levels (a) and (b) better than 70%. However, the performance drops off for the higher energy consumption levels' samples, being the same or worse than for pitch. This provides an indication of the existence of a tempo-urgency threshold, beyond which users find all cues to indicate extreme energy usage and thus struggle to make consistent associations. Accordingly some further method of differentiation may be required to make higher tempo sounds more distinguishable from one another.

Accuracy: The standard deviation and offset plots show that the tempo cues performed better than the pitch cues on both the bottom and top ends of the scale, especially the latter. This indicates that increasing urgency can be imparted via the use of higher tempos, and that they also perform well in the lower range. Given that all of the tones used in this test were of the same pitch, and thus could be reproduced with equal presence through all varieties of audio-speaker, it can be concluded that the use of tempo is likely considerably more effective for expressing lower urgency levels than pitch would be under good sonic conditions.

Tempo and Texture

Consistency: The introduction of the texture parameter significantly improves the consistency at level (c) and (e) by 10% and 15% respectively, at the expense of level (b), down by 25%. This points to the conclusion that a full five levels of urgency may not be practical for an audio interface of this nature. Rather, the use of a maximum of three notification levels would likely lead to better results, with users making the correct associations far more easily. This category is the most consistent for the extreme low and high levels.

Accuracy: Levels (a) and (e) also show a small offset (<0.25) and a low standard deviation (<0.5). The offset and standard deviation of levels (c) and (d) remain unchanged when adding texture. This category is also the most accurate for the extreme low and high levels.

Animal Themes

Consistency: This set of audio cues yielded the highest consistent results overall (>70%), indicating that respondents found it easy to make associations between the sounds and energy consumption levels.

Accuracy: The mean and standard deviation of the offset are the worst of the four categories. These sounds contain significant meaning making them easy to distinguish, but indicates that each user interprets the sound differently. If similarly complex sounds samples can be found that generate more universal associations within users, then the approach could yield far more accurate results.

The majority of respondents indicated that they would take physical action to reduce energy consumption around the audio cue level they had perceived to be associated with level (c) usage. Respondents may have interpreted this question to be an assessment of their own commitment to energy reduction, and may thus have chosen a moderate response level that they felt to be the appropriate response. However, the responses tail off in both the high and low directions, providing at least some basic indication that the overall range of urgency covered in the tests is centered around a level where an active user response may be triggered.

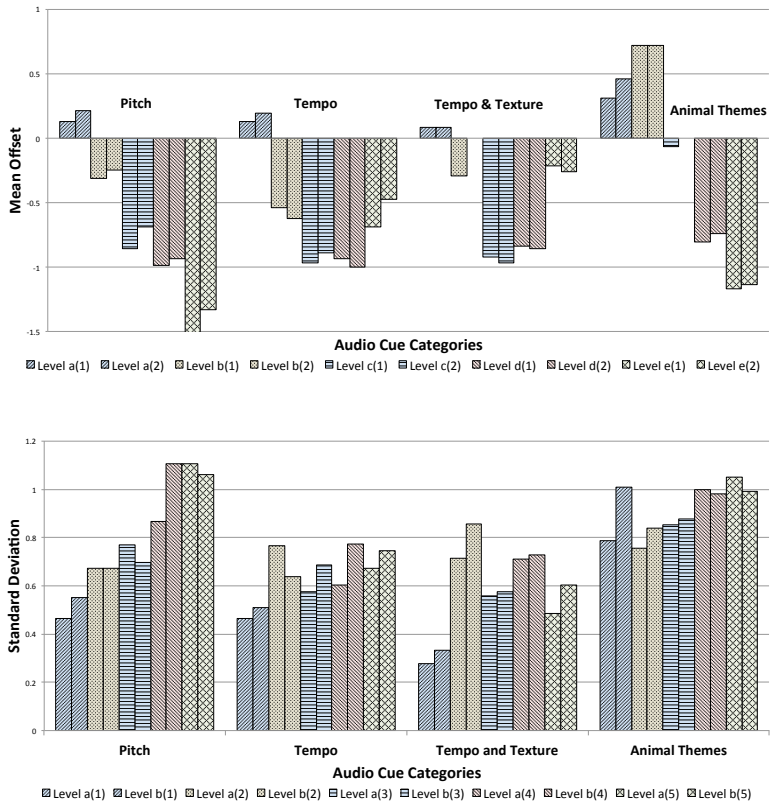

Fig. 3. The mean and standard deviation of the offset between actual and perceived level for second and third instance of each audio cue

6.2 Criterion Three

This criterion states that the feedback provokes a response at the appropriate point in time. As this test was not conducted in a live setting, users where instead asked at which level they would consider taking action. The results are shown in Fig. 4.

Reduction in energy usage is a desirable outcome at any level of consumption. Indeed a small number of people (3%) stated that would take action at level (a), and 30% at level (b). A significant number (77%) stated they would take action at level (c). This result gives an optimistic outlook that there is a fair chance that this would happen in practice, but this cannot be conclusively stated at this stage of the research.

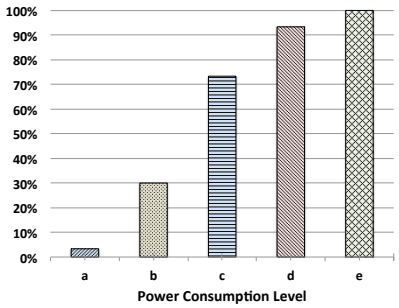

Fig. 4. Cumulative user perception of level at which action should be taken to reduce consumption

6.3 Criterion Four

This criterion states that the feedback does not overburden or confuse the user. This was assessed by analysing the responses to the question "Please tell us which set of three tests you thought were the most effective and why." and shown in Fig.5. This data indicates the users perception of the efficacy of the different categories of audio cues.

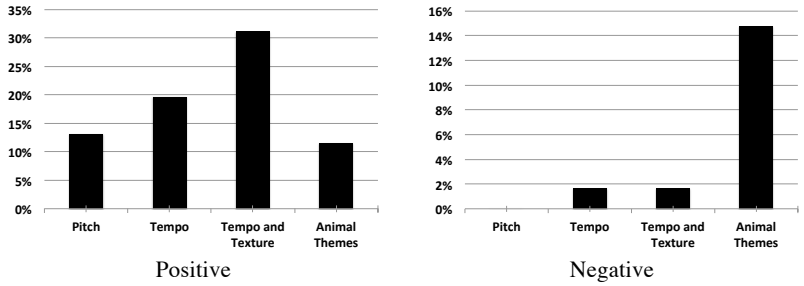

Fig. 5. User perceptions of test effectiveness by audio cue category

The combined tempo and texture category was perceived to be most effective (>30%). This result tallies with the analysis presented in sub-section 6.1. Respondents deemed the standalone tempo cues to be the next most effective, although considerably less so, receiving 37% less positive feedback than the combined cues. The standalone pitch and animal sound variants of the test were found to be the least effective, receiving 58% and 63% less positive feedback respectively than the combined tempo and texture test.

The overwhelming majority of negative perceptions were aimed towards the animal themed audio cues, users finding them to be either high in annoyance factor or challenging to rank in terms of urgency (again backed up by the numerical analysis). It should be noted that many respondents acknowledged the tempo element of the tempo and texture audio cues as being a significant contributor to that test's effectiveness. Thus, when considered in combination with the positive feedback recorded for the tempo and texture test, respondents can be considered to have found tempo to be the most effective parameter by a wide margin.

Whilst this study has tested for four specific audio traits, it would be of considerable interest to investigate a wider range of properties, such as rhythm, melody and harmony, as well as testing further cultural associations beyond animal themes. However, these would have to be applied in such a manner that they also featured low irritation indices for lower energy usage levels, perhaps via combined usage with tempo, volume or frequency of performance.

7 Conclusion

This paper addresses the issue of developing a design methodology for providing immediate and intuitive audio feedback about high power consumption, especially during periods when the power level is approaching the capacity of the main circuit breaker.

The four criteria used in this study for the assessment of the efficacy of the feedback mechanism are that: it is appropriate to the specific context in which it is intended to be used; it is interpreted consistently and accurately, provokes a response at the appropriate point in time, and does not overburden or confuse the user.

Due to the specific requirement for an immediate response is, independent of the location of the user, and that multiple levels of feedback are useful, a group of five audio cues were used. Four categories of cues were developed - three based on fundamental musical properties: pitch, tempo and tempo-with-texture, and one based on complex sounds (animal noises).

A survey of 61 respondents showed that the tempo-with-texture category best met the four requirements.

References

1. Paton, C.: Eskom Was 'On the Brink of a Power Shutdown'. Business Day Live, May 20, 2003. Retrieved February 1, 2014, http://www.bdlive.co.za/business/energy/2013/05/20/eskom-was-on-the-brink-of-a-power-shutdown (2003)
2. Zehir, M.A., Bagriyanik, M.: Demand Side Management by Controlling Refrigerators and its Effects on Consumers, Energy Conversion and Management, 64 (1). pp. 238–244 (2012)
3. Van Blommestein, K.C., Daim, T.U. Residential Energy Efficient Device Adoption in South Africa. Sustainable Energy Technologies and Assessments, 1 (1), pp. 13–27 (2013)
4. Xia, X., Setlhaolo, D., Zhang, J.: Residential Demand Response Strategies for South Africa. In: IEEE Power and Energy Society Conference and Exposition in Africa (PowerAfrica), pp. 1–6. IEEE (2012)
5. Tewari, D., Shah, T.: An Assessment of South African Prepaid Electricity Experiment, Lessons Learned, and Their Policy Implications for Developing Countries. Energy Policy, 31 (9). pp. 911–927 (2003)
6. CBi-electric Load Control Relay. Retrieved February 3, 2014, http://www.cbi-electric.co.za/products_select.php?p=13#245 (2014)
7. He, H. A., Greenberg, S., Huang, E. M.: One Size Does Not Fit All: Applying the Transtheoretical Model to Energy Feedback Technology Design. In: Proc. CHI'10, pp. 927-936. ACM Press (2010)
8. Allcott, H., Mullainathan, S.: Behavior and Energy Policy. Science, 327 (3). pp. 1204-1205 (2010)
9. Ehrhardt-Martinez, K., Donnelly K. A., Laitner, J. A.: Advanced Metering Initiatives and Residential Feedback Programs: a Meta-Review for Household Electricity-Saving Opportunities. American Council for an Energy-Efficient Economy, Washington (2010)
10. Froehlich, J., Findlater, L., Landay, J.: The Design of Eco-Feedback Technology. In: Proc. CHI'10. pp. 1999-2008. ACM Press (2010)
11. Fitzpatrick, G., Smith, G.: Technology-Enabled Feedback on Domestic Energy Consumption: Articulating a set of design concerns. IEEE Pervasive Computing, 8 (1), pp. 37-44. IEEE (2009)
12. Ford, R., Karlin, B.: Graphical Displays in Eco-Feedback: A Cognitive Approach. In: Proc. HCII'13. Springer (2013)
13. Util Labs Low Voltage Smart System. Retrieved February 4, 2014, http://www.utillabs.com/sites/default/files/news-events-downloads/lvss_brochure_final_web_2.pdf (2011)
14. Walker, B. N.: Consistency of magnitude estimations with conceptual data dimensions used for sonification. Applied Cognitive Psychology, 21 (5) pp. 579–599, John Wiley & Sons (2007)
15. Fogg, B. J.: A Behavior Model for Persuasive Design. In: Proc. International Conf. on Persuasive Technology, pp. 40-46, ACM Press (2009)
16. Kallinen, K.: Emotional Responses to Single Voice Melodies: Implications for Mobile Ringtones. In: Human-Computer Interaction - Interact, pp. 797-800 (2003)