

MODELLING OF MODULAR BATTERY SYSTEMS UNDER CELL CAPACITY VARIATION AND DEGRADATION

DANIEL J. ROGERS

Department of Engineering Science, University of Oxford, UK

LOUIS J. M. ASLETT

Department of Mathematical Sciences, Durham University, UK

MATTHIAS C. M. TROFFAES

Department of Mathematical Sciences, Durham University, UK

ABSTRACT. We propose a simple statistical model of electrochemical cell degradation based on the general characteristics observed in previous large-scale experimental studies of cell degradation. This model is used to statistically explore the behaviour and lifetime performance of battery systems where the cells are organised into modules that are controlled semi-independently. Intuitively, such systems should offer improved reliability and energy availability compared to monolithic systems as the system ages and cells degrade and fail. To validate this intuition, this paper explores the capacity evolution of populations of systems composed of random populations of cells. This approach allows the probability that a given system design meets a given lifetime specification to be calculated. A cost model that includes the effect of uncertainty in degradation behaviour is introduced and used to explore the cost-benefit trade-offs arising from the interaction of degradation and module size. Case studies of an electric vehicle battery pack and a grid-connected energy storage system are used to demonstrate the use of the model to find lifetime cost-optimum designs. It is observed that breaking a battery energy storage system up into smaller modules can lead to large increases in accessible system capacity and may lead to a decision to use lower-quality, lower-cost cells in a cost-optimum system.

1. INTRODUCTION

For electric vehicles and grid-connected energy storage systems, the construction of large, reliable battery packs is a crucial technological hurdle. This is because the quantity of stored energy required in these applications is far in excess of that which can be provided by a single cell: compare for example a 10 Wh 3.7 V single-cell

E-mail addresses: dan.rogers@eng.ox.ac.uk, louis.aslett@durham.ac.uk, matthias.troffaes@durham.ac.uk.

mobile phone battery to a 50 kWh 400V PEV battery pack containing 400 or more cells, or to a 10 MWh grid storage system connected at medium voltage that may contain 100,000 cells. In these large systems, battery packs are formed by connecting cells in series in order to provide a pack terminal voltage that is suited to the application. An individual electrochemical cell is a complicated object consisting of tens or hundreds of layers of anode, cathode, electrolyte and separator materials which exhibit complex electrochemical, thermal and mechanical behaviour. Differences between cells stemming from manufacturing tolerances and variations in operating conditions should therefore be expected. As a result, although the series connection process inherent in forming a battery pack is conceptually simple, it gives rise to a multitude of engineering challenges attached to the management of many similar-but-not-identical cells that must perform predictably and repeatably in the context of an overall energy storage system.

Although manufacturers take great care to deliver uniform cell populations to their customers, it has been shown numerous times that even nominally identical cells using mature technology and taken from the same batch demonstrate significant variations in initial capacity [16] and other parameters such as impedance [23]. It is also well-known that cells display varying degradation rates dependent on discharge rate [11] and ambient temperature [12], for example. The authors of long-term cycling studies comment on the unpredictability of the cell-to-cell behaviour, in the sense that the capacity observed during cycle N is not a particularly strong predictor of, for example, capacity at cycle $N + 200$, and that the rate of capacity loss varies over time such that cell capacities “*cross each other*” [15] during the lifetime of the system: the highest capacity cell in a batch at time T is not particularly likely to be so at time $2T$.

In this work, we focus on the interaction between uncertain cell degradation behaviour and Energy Storage System (ESS) design because we believe statistical modelling of large ESSs has been under-explored in the literature to-date and that there are useful lessons to be learnt by abstracting the problem away from particular chemistries and particular applications. The problem is approached by defining a simple statistical cell model that qualitatively replicates the behaviour seen in previous experimental studies of cell degradation. The model lends itself to studying the behaviour of large ESSs through analytical techniques and large-scale numerical simulation.

Underlying this study is the general concept of ESS *modularisation*: this is the process of breaking an ESS up into sub-units (modules) so that differences between modules do not affect operation of the other modules. In this way, cell-to-cell variation, cell degradation and cell failure is acceptable because although this may cause a module to perform poorly or fail, the ESS (composed of many modules) can continue to function, albeit with a small overall reduction in performance. The engineering justification for modularisation is clear: for a given cell type and application, there will be an upper limit on how many cells can be connected together in series before the probability of failure of the string becomes unacceptably high. This general idea is sometimes explicitly, but often implicitly, discussed in the ESS literature. In [6] three different designs for a 1 MWh grid-connected ESS were compared, all of which were modularised in some fashion in order to provide adequate reliability (cells were assumed to have a constant or temperature-dependent failure rate). In [18] several pack configurations for a grid-connected ESS were studied,

with the authors noting “*it is not recommended to let the number of cells in series go over 200. An appropriate tradeoff needs to be made between the battery module voltage level and the system reliability.*” A probabilistic approach is used to compare different EV battery pack layouts in [19], and the authors show that although the prediction of pack state-of-health becomes less accurate as the packs age, this can be mitigated to some extent through appropriate design of the pack. There are many studies that present what are, in-effect, methods to achieve modularisation, although often the authors do not describe it in this way; terms such as ‘cell balancing’ [13, 21, 7, 5], ‘cell equalization’ [10] or ‘pack reconfiguration’ [20] may be used instead.

It is important to draw a distinction between *modularisation* and *cell balancing*: a system that implements a cell balancing mechanism is able to charge and discharge different cells at different rates in order to accommodate cell-to-cell variation in capacity. The differential rate capability (i.e. how different the charging/discharging rates can be made from one cell to another) may be relatively small, perhaps 5-20% of nominal rates. In such cases, the system can be considered to be modularised in the way meant in this work, but this modularisation only works so long as the differential-rate capability is not exceeded (i.e. the cells do not develop excessive variation). Specifically, in a simple series-connected pack (which often contain differential-rate limited cell balancing circuits of some form or another) it is normally not possible to isolate a completely failed cell, but in some more-complex designs it is possible [6, 5, 20]. This work studies hypothetical systems that are truly modularised, i.e. ones that can bypass cells completely, in order to avoid the need to explicitly choose a somewhat arbitrary differential-rate capability. However, many of the key observations apply to systems using balancing-type circuits.

In this study, we use Monte Carlo simulation [22] to quantify the effects of variability in cell capacities and in the degradation process when designing modularised battery systems, following a similar approach to simulation in reliability analysis of power systems [2, 3]. To see how these effects impact design decisions, we use standard statistical decision analysis [8, 17].

The paper is structured as follows: Sections 2 and 3 introduce a generic statistical model of a cell, a definition of the process of modularisation and presents some general results. Section 4 provides a numerical investigation of the behaviour of a population of cells operating in a modular battery system. Section 5 introduces cost and revenue models which are used to explore optimal design of modular battery systems. Section 6 concludes the paper and highlights the contributions of the work.

2. SYSTEM MODELLING

The *capacity* $C_i(t)$ of a cell (where i denotes the cell index, and t denotes time) is defined as the total energy that the cell will release when taken from a fully-charged to fully-discharged state. The cell capacity is an indicator of the remaining economic value of the cell; once cell capacity is zero the cell performs no useful energy storage function. A pack-producer is likely to have quite precise knowledge of initial capacity $C_i(0)$, as this is easily measured. As time t increases, the cell will degrade, and thus $C_i(t)$ will be a decreasing function. The degradation rate is typically uncertain, and will vary between cells.

The *accessible capacity* $C(t)$ of a system is defined as the total energy that the system will release when taken from a fully-charged to fully-discharged state. In the ideal case the accessible capacity will be equal to the sum of all the capacities of the cells it contains:

$$(1) \quad C_{\text{ideal}}(t) := \sum_{i=1}^n C_i(t)$$

However, in practical systems, the accessible capacity will almost always be smaller than the underlying sum of all individual cell capacities. In a simple series-connected string of cells, each cell can deliver only the capacity equal to the capacity of the lowest capacity cell (this is because the (dis)charge process must stop when any one cell is fully (dis)charged, as attempting further (dis)charging will damage the cell permanently):

$$(2) \quad C_{\text{series}}(t) := n \min_{i=1}^n C_i(t)$$

Clearly, $C_{\text{series}}(t) \leq C_{\text{ideal}}(t)$ where equality obtains when all cells have identical capacities. The inefficiency increases as cell capacities become less homogeneous.

2.1. Modularisation. In order to mitigate this effect, and as discussed previously, larger battery stems are often modularised in some way. In this work, this is taken to mean that m series strings of cells, each containing ℓ cells, are combined into a larger system such that all the accessible capacity of each string can be delivered by the system. The accessible system capacity of a modularised system is

$$(3) \quad C(t; \ell, m) = \sum_{i=1}^m \ell \min_{j=1}^{\ell} C_{ij}(t) = \sum_{i=1}^m \ell C'_{i1}(t)$$

where $C'_{ij}(t)$ denotes the j th smallest element of $C_{i1}(t), \dots, C_{i\ell}(t)$.

For $m = 1$ ($\ell = n$), Equation (3) is equivalent to Equation (2) and the system is composed of a single large series string; in this case the system is not modularised. For $m = n$ ($\ell = 1$), Equation (3) is equivalent to Equation (1) and the capacity of every cell is fully accessible; in this case the system is fully modularised. Clearly, for a fixed number of cells $n = \ell m$, as m increases, the accessible capacity is likely to increase because a low-capacity cell in one module does not affect other modules. As cells age, their capacities will tend to diverge and so a highly modularised system will tend to maintain larger accessible capacity over time. After failure of one cell in a string, the accessible capacity of the string containing that cell will be zero and so, if cell failure rate is relatively high compared to the intended lifetime of the system, modularisation is necessary to maintain functionality for a longer period.

2.2. System metrics. A useful metric is the *Accessible Capacity Fraction* (ACF):

$$(4) \quad \text{ACF}(t; \ell, m) := \frac{C(t; \ell, m)}{\sum_{i=1}^n C_i(t)} \in [0, 1]$$

This represents the fraction of underlying cell capacity that is accessible in a particular system design at any point t in the life of that system, i.e. how close to complete utilisation of its cells a system gets. $\text{ACF} = 1$ if $\ell = 1$ and in general $\text{ACF} < 1$ if $\ell > 1$ unless all cells are perfectly matched. ACF will tend to fall over

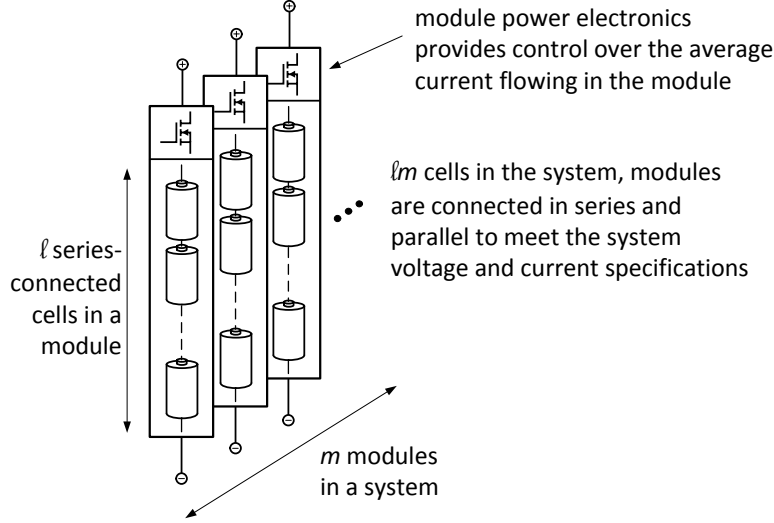


FIGURE 1. Conceptualisation of a modularised battery pack. Cells inside a module are connected in series and have equal currents. The participation of a module in the overall system is controlled by the module power electronics. The practical realisation of the power electronics will vary depending on the application.

time as cells become less well matched. A second useful metric is the *Accessible Integrated Capacity Fraction* (AICF):

$$(5) \quad \text{AICF}(T; \ell, m) := \frac{1}{T} \int_0^T \text{ACF}(t; \ell, m) dt \in [0, 1]$$

The AICF provides an indication of the advantage of modularisation accrued over the lifetime of the system. An AICF close to 1 indicates the design is able to exploit the capacities of the cells it contains over the life of the system.

2.3. Optimal ordering of cells. An intuitive idea is that, during manufacture, the accessible capacity of a new system may be maximised by grouping similar-capacity cells together in modules because this leaves the least capacity inaccessible in each module. Formally, it is shown in Appendix A that $C(t; \ell, m)$ (and therefore the ACF and AICF) is maximal when the ordering of the cell capacities matches the lexical ordering of the cell indices, i.e. when the ordering of $C_{ij}(t)$ matches the natural ordering of $\ell(i-1) + j$. This also applies at any point t in the lifetime of the system, i.e. there is an optimal rearrangement of cells which is made by swapping cells between modules that results in the highest accessible system capacity. However, in practice, this is likely to be a highly invasive procedure which may be cost-prohibitive.

3. CELL MODELLING

3.1. A Simple Degradation Model. In this work, we use the term *cell degradation* to simply mean a decrease in cell capacity over time. In practice, cell degradation is a complex mixture of chemical and mechanical processes that vary

from one cell type to another, not all of which are fully understood [4]. All cell types tend to experience a loss of active material over time due to slow but irreversible side-reactions occurring within them, even when the cell is not actively cycled. For most lithium ion cell types, the intercalation of ions in porous anode and cathode materials during cycling produces mechanical stresses that cause gradual accumulated damage to the structure of the cell, leading to reduced capacity. Degradation rates tend to increase if a cell is operated at particularly high or low temperatures and/or at high charge or discharge currents, i.e. cell degradation depends on how the cell is operated. We do not attempt to model any one particular degradation mechanism or group of mechanisms, but instead to relate the general expected degradation behaviour of a particular set of cells to the overall behaviour of a system composed of these cells.

The underlying requirement is to construct $C_i(t)$ such that it is a non-increasing random process that qualitatively replicates the degradation behaviour that has been observed experimentally in previous studies (specifically the variation of degradation in a population of similar cells). We note that articles containing information of the sort required are relatively rare, likely because of the large-scale facilities and long timescales required to test many similar cells over hundreds of cycles. The following three studies have, however, performed this type of work and therefore help provide a justification for the design of the model:

- (1) [1] describes an experiment on 48 1.85 Ah carbon/nickel manganese oxide (NMC) cylindrical cells operated for 1700 charge/discharge cycles under the same conditions for all cells (1.08C charge and discharge). The authors observed a clear distinction between “*two ageing mechanisms with different degradation rates*” in a plot of cell capacity versus cycle number. This plot exhibits a breakpoint at approximately 1100 cycles and 10-15 % drop in capacity. Before this breakpoint, cell capacity fell on average at a rate of 0.017 % capacity loss per cycle, and after at 0.063 % per cycle.
- (2) [15] describes an experiment on 24 4.4 Ah graphite/lithium cobalt oxide (LCO) pouch cells operated for 593 cycles (1C charge, 10C discharge). Most cells display a breakpoint in the degradation data, with slower degradation (0.04 %/cycle) occurring before and faster degradation (0.12 %/cycle) afterwards. However, cell behaviour is different from that observed in [1] in that the breakpoints tend to occur earlier and are more widely distributed in proportion to the length of the experiment (from 180 cycles onwards), although the capacity drop at which it occurs is similar at 8-15 %.
- (3) [24] describes an experiment on 124 1.1Ah graphite/lithium iron phosphate cells (LFP) operated for 1000 cycles. A major difference versus [1, 15] is that the charging conditions were not the same for all cells (charging varied between 3.6 C and 6 C, discharging was 1 C for all cells), leading to very different degradation rates across the population. However, a similar breakpoint effect is observed for most cells, with the cycle number at which the breakpoint occurs being lower for higher charging rates. The breakpoint tends to occur after 400 cycles and 4-10 % capacity drop, although there are three outlier cells that degrade very rapidly and several cells which appear to not reach the breakpoint during the experiment.

Although only presenting results for three lithium titanate (LTO) cells, [14] explicitly fits a “*two-stage piecewise linear model*” of degradation to experimental

data. In [25] cell resistance and power delivery capability exhibit a breakpoint, suggesting that degradation mechanisms affect other cell parameters beyond capacity. There are studies where a breakpoint in degradation rate is not observed, for example [9] presents results for 51 18650-format LCO cells showing constant cell degradation rates at 0.25 C and 1.5 C rates over 1000 cycles, although there is significant variation in this rate between cells.

Based on these observations, we propose the following simple degradation model for the capacity of an individual cell:

$$(6) \quad C_i(t) = \begin{cases} \max\{0, C_i(0) - D_i t\} & \text{if } t < T_i \\ \max\{0, C_i(0) - D_i t - E_i(t - T_i)\} & \text{if } t \geq T_i \end{cases}$$

where $C_i(t)$ is the capacity of cell i at time t , D_i is the degradation rate, $D_i + E_i$ is the accelerated degradation rate, and T_i is the breakpoint time of transition to accelerated degradation. Random variables are denoted by capital letters, and their realisations by lower case letters. It is assumed that $C_i(0)$, D_i , E_i , and T_i are independent and follow non-negatively truncated Normal distributions. This model conforms to the behaviour of lithium-ion cells reported in the literature and by application engineers: a relatively narrow spread in initial capacity that diverges due to degradation until capacity has dropped by 5-15%, followed by more rapid degradation and divergence until the cell is deemed to have failed. The definition of cell failure is application-specific; it could be related directly to capacity (i.e. because the capacity of the cell is now inadequate to serve the intended role) or related to its electrical performance (for example, internal impedance increase leading to inadequate charge/discharge rate capability). We do not attempt to model any failure mechanism beyond capacity degradation, although this could be added to the model if sufficient data were available to support it. Tailored models could of course be built for a particular cell design by performing large-scale tests, with functional relationships and parameters statistically fitted and validated accordingly. However, the model as proposed exhibits the main features of typical cell behaviour and allows us to demonstrate a methodology for studying and quantifying the impact of cell variability on pack design and eventual system profitability.

3.2. Limitations of the model. An attraction of the model is that it is independent of engineering implementation detail and is therefore relatively simple and capable of providing general insight. However, the assumption of independence between cells is potentially naïve. Rates of degradation may be positively correlated between nearby cells in a pack if, for example, the thermal management system is inadequate and leads to hotspots forming in the pack. A major cell failure in a practical system could occasionally result in thermal runaway of the cell and, in the worst case, a cascading failure. Such correlations could not be modelled without significant engineering detail of the proposed system and so these are not included here. This means that to some degree, the model relies on the observation that a well engineered system should not contain obvious weaknesses (e.g. by allowing hotspots to form); it appears reasonable to assume that correlated capacity loss (and/or outright failure) is avoided as a result of sensible design and maintenance practices. Unusual degradation or frequent failures would lead to the system being improved upon until these were avoided. As a result, the framework presented here

| Parameter | Good cell | Bad cell |
|----------------|-----------|----------|
| σ_{C_0} | 0.01 | 0.03 |
| σ_D | 0.02 | 0.05 |
| σ_T | 0.1 | 0.2 |
| σ_E | 0.1 | 0.2 |

TABLE 1. Parameter values used for numerical simulation study, representing typical behaviour of ‘good’ and ‘bad’ cells. Both ‘good’ and ‘bad’ cells have means $\mu_{C_0} = 1, \mu_D = 0.2, \mu_T = 1$ and $\mu_E = 0.6$.

represents a well designed pack that is only affected by unavoidable and unpredictable cell degradation behaviour.

Variation in cell behaviour is dictated solely by the parameters given in Table 1. The model does not include an equivalent circuit or electrochemical submodel. As a result, instantaneous dependence of degradation rate (e.g. higher degradation rates at higher C-rates) or the effect of operating conditions over the lifetime of the system (e.g. seasonal ambient temperatures) are not explicitly modelled. Such effects could be approximated by modifying the cell parameters to reflect the anticipated use case, i.e. the degradation parameters in Table 1 could be increased if the system was to be operated at a high discharge rate or under cold conditions, for example.

4. NUMERICAL SIMULATION OF SYSTEM BEHAVIOUR

The cell model is now used to perform a set of numerical simulations of systems composed of two types of cells, the parameters of which are given in Table 1. These cells are labelled ‘good’ and ‘bad’. Good cells are characterised by tight statistical distributions such that the capacities of all cells start close and stay relatively similar over time across the system. Conversely, bad cells are characterised by loose statistical distributions such that their capacities are relatively dissimilar and tend to diverge quickly. The good cells represent the output of a well controlled, close-tolerance manufacturing process, perhaps with a high fraction of rejects, and are likely to be relatively expensive. The bad cells represent the opposite: a relatively low-tolerance manufacturing process that does not reject a large number of cells. Of course, it is likely that the bad cells will cost less than the good cells, and it is perfectly feasible that a well-designed system composed of bad cells will be capable of serving the same application as a system composed of good cells, but at a lower cost.

The following sections use the ACF and AICF metrics to quantify the value of systems composed of good and bad cells. We seek to answer the question: is it better to build battery systems using relatively few, large modules made up of expensive, tightly distributed good cells, or use many, small modules made up of cheaper bad cells?

The average starting capacity ($\mu_{C_0} = 1$), average initial degradation rate ($\mu_D = 0.2$), the average breakpoint time ($\mu_T = 1$) and the average accelerated degradation rate ($\mu_E = 0.6$) are the same for good and bad cells. This means that, for both good and bad cells, the mean cell starts with a capacity of 1, degrades to a capacity of 0.8 in time 1 and then degrades to zero capacity in a further time 1. Good and bad cells differ only in the tightness of the distribution of the parameters that

control degradation, not their mean values: as a result, the expected capacity of good and bad cells are equal at any time instant.

Figure 2 depicts the behaviour of modularised battery storage systems. The system is composed of 100,000 cells and is large enough to demonstrate the important statistical behaviour of the system. It was observed that initial optimal ordering of cells (as per Theorem 1, Appendix A) improves ACF by roughly 1-5% in most cases, but does not qualitatively affect the overall pattern of the results (of course, in the case of $\ell = 1$, the ordering of cells is irrelevant because every cell has its own module). In the following discussion we assume this initial ordering is conducted by the system manufacturer in order to gain this improvement in system performance at what is likely to be a relatively low cost.

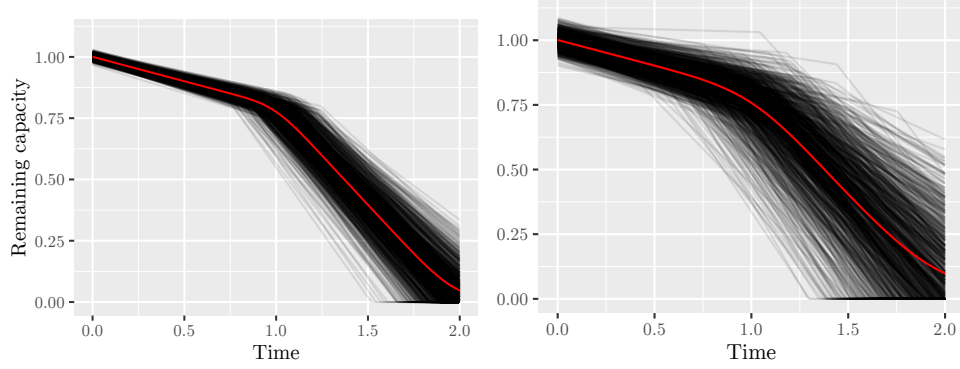
More highly modularised systems (i.e. smaller ℓ) provide a greater ACF, particularly once a significant fraction of cells begin accelerated degradation at approximately times 0.8 (good cells) and 0.4 (bad cells). In addition, highly modularised systems provide significantly longer periods of time with high ACF. For example, a good-cell system with $\ell = 10$ provides $\text{ACF} > 0.75$ up to time 1.4 (versus 1.0 for $\ell = 160$), and a bad-cell $\ell = 10$ system provides $\text{ACF} > 0.75$ up to time 1.15 (versus 0.75 for $\ell = 180$).

The AICF plots emphasize the cumulative effect of modularisation: highly modular systems benefit from higher ACF for longer. In practice, it is unlikely that any system would be operated until $T = 2$ as at this time all cells are severely degraded and so the system would only deliver a tiny fraction of its original capacity, likely making operation economically unviable. Instead, system end-of-life may be considered to be $T = 1$. At this time, half the cells have experienced the accelerated degradation breakpoint and so average cell capacity begins to fall rapidly after this time. By $T = 1$, a good-cell $\ell = 10$ system will deliver approximately 6.5% more AICF than a $\ell = 10000$ system. For a bad-cell system this increase is 31%, demonstrating that a high degree of modularisation is particularly valuable for systems where the spread in cell degradation is not well-controlled.

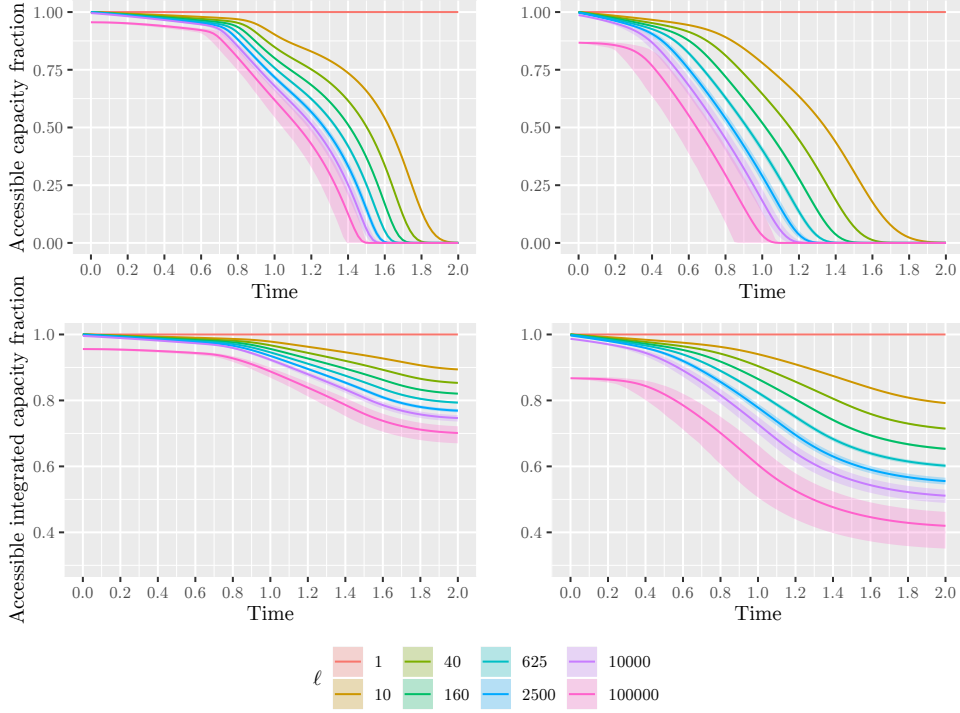
4.1. Other parameter selections. The cell parameters in Table 1 could instead be chosen to compare the performance of different cell chemistries in a particular application. For example, when designing an ESS for a high-temperature environment, the designer may wish to compare a ‘standard’ cell chemistry with relatively high degradation rates at elevated temperatures, against a ‘specialised’ cell chemistry that has lower energy density but also lower degradation rates. The specialised cell would be modelled with a lower μ_{C_0} but also lower μ_D and μ_E and higher μ_T . In another scenario, a designer might wish to compare two ESS cooling systems, one which ensures very even temperature between cells and another, cheaper, solution that allows more temperature variation and so causes a wider spread in degradation rates: this could be modelled by increasing σ_D , σ_E and/or σ_T for the cheaper design.

5. OPTIMAL SYSTEM DESIGN

5.1. System cost. Modularisation comes at a cost in terms of additional power electronics, sensing and communication hardware required to implement each module. The total cost of a modularised system can be modelled as a cost k_1 per cell, a fixed cost k_2 per module and a module cost k_3 that scales with the size of the



(A) Random realisations of capacities of good (left) and bad (right) cells, with mean shown in red.



(B) ACF and AICF of good (left) and bad (right) cells, for different levels of modularisation in a system with 100,000 cells.

FIGURE 2. Comparison of modularised systems composed of good or bad cells. Coloured bands show 95% confidence intervals.

module. k_2 represents the one-off module costs associated with sensing and communication electronics. k_3 represents the costs that scale with the size of the module, for example mechanical housings, wiring, and the requirement for the power electronics to handle a power proportional to module size. This leads to a total system

cost:

$$(7) \quad K(\ell, m) := m(k_1\ell + [k_2 + \ell k_3])$$

Setting $k_1 + k_3 = 1$ has the effect of normalising the cost equation to the cost component that scales linearly with system cell count ($m\ell$), i.e. the marginal cost of adding a cell to the system. This gives system cost:

$$(8) \quad K(\ell, m) := m(\ell + k)$$

where k represents the cost, relative to the marginal cost of adding a cell to the system, of the one-off hardware cost to form a new module. The costs presented in the numerical studies that follow involve only one free cost parameter k and can be interpreted as being normalised to the cell marginal cost.

5.2. System profit. The value (profit) of the system is its ability to deliver an intended service over a period of time. Practically, many external factors are likely to influence profit; for example both revenue and costs are likely to be dependent on the evolution of market conditions over the lifetime of the system. Here, we ignore external factors and explore a simple approach for linking system ACF to a quantified profit, thereby allowing the value of modularisation to be explored.

5.3. Lifetime capacity profit model. This represents a scenario where a system operator is paid per unit of energy exchanged with a supply/load over the lifetime of the system. This is a good representation of the revenue model for a grid-connected arbitrage application, for example. It is assumed that the system is fully utilised and in continuous operation so that it exchanges its full capacity with the grid many times over its lifetime. In this case, the instantaneous value of the system is proportional to its capacity, i.e. the revenue to be generated by one full charge/discharge cycle performed at a particular moment in the life of the system. The accumulated lifetime revenue is

$$(9) \quad R_1(\ell, m) := \alpha_1 \int_0^T C(t; \ell, m) dt$$

where α_1 is a proportionality constant that places a value on the time-capacity product delivered by the system.

The goal of the system designer is to maximise the expected profit by choosing the optimal level of modularisation, given the cost of modularisation relative to the marginal cost of the cells (k), for a fixed number of cells in the system (n):

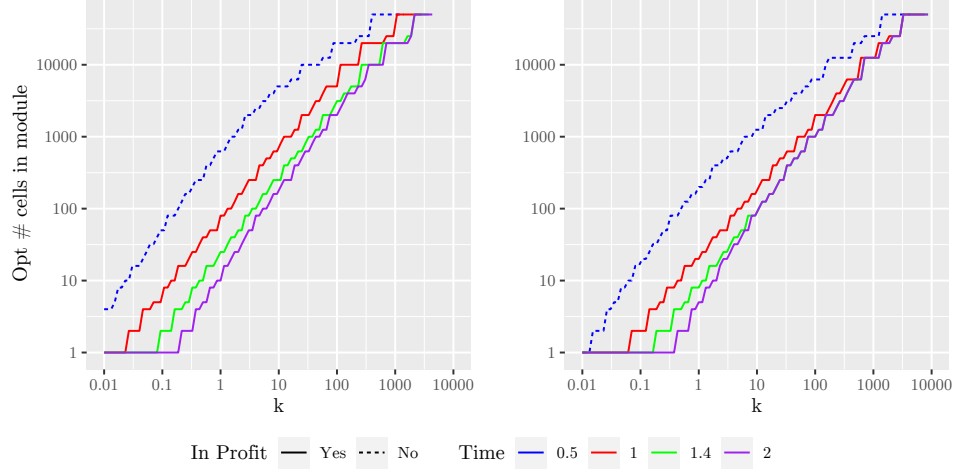
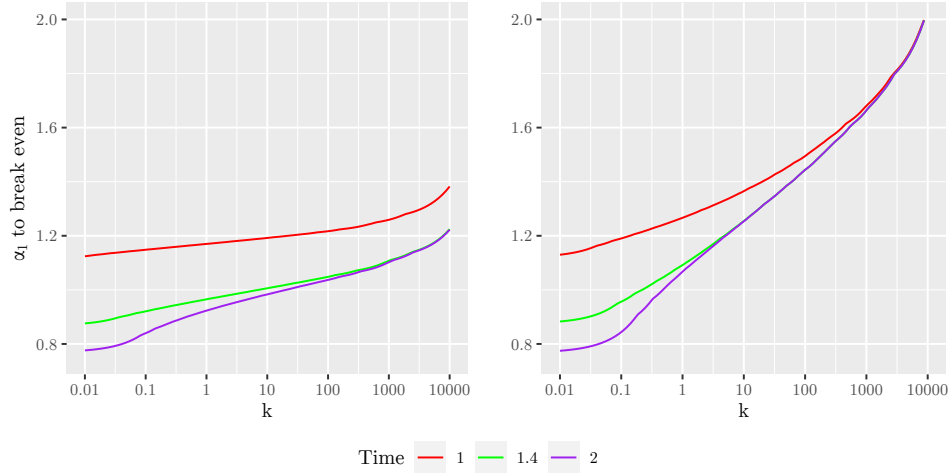
$$(10) \quad P_1(\ell, m) := \mathbb{E}[R_1(\ell, m)] - K(\ell, m)$$

subject to

$$(11) \quad \ell m = n$$

This is an integer programming problem as ℓ and m must be positive integers. Figure 3 shows the result of this optimisation with $\alpha_1 = 2$, $n = 100000$ as k is varied between 10^{-2} and 10^2 . When the cost of modularisation is low, smaller module sizes are favoured because of the increase in AICF that this enables. Smaller module sizes are also favoured if the system is allowed to run for longer periods of time because small modules are better able to tolerate the wide variation in cell capacities that occurs near end-of-life.

It is instructive to study the value of α_1 that results in break-even cost (i.e. $P_1(\ell, m) = 0$) for the optimal ℓ . This indicates the revenue required per-cell in

(A) Optimal module size for good (left) and bad (right) cells, with $\alpha_1 = 2$.(B) α_1 required for break-even for good (left) and bad (right) cellsFIGURE 3. Optimal design of systems composed of 100,000 good or bad cells as function of k , the ratio of the incremental cost of a cell to the cost of modularisation.

order for the optimal system to break-even. One way of thinking about this is to imagine a system composed of perfectly matched, non-degrading cells (at a marginal cost of 1 each) that are assembled into a system that has no other costs: effectively a system made of one large module ($\ell = n$), such that the one-off module costs are negligible and so $K = n$. If this system were operated for time 1, the revenue would be $R_1 = n$ and the break-even α_1 would be 1. Lower break-even α_1 will be achieved if the system is operated for a longer time, higher break-even α_1 will occur if modularisation electronics is required (and $k > 0$) or cells degrade.

Figure 3 shows three predictable effects: Break-even α_1 is lower when the system is run for longer periods of time, is lower when the cost of modularisation is lower, and is lower when using good cells. Perhaps less obvious is the relatively weak effect of the cost of modularisation on break-even α_1 : For example, when using good cells, large changes to the cost of modularisation (k changing from 10^{-2} and 10^2) increases break-even cost by a relatively small amount ($< 8\%$). This is because module sizes remain large and so the total additional cost attributable to modularisation is low. A system constructed using bad cells tends to use smaller modules and so the gradient of the break-even α_1 line is steeper. An interesting result is that, especially for long system lifetimes and low modularisation costs, the bad cell break-even α_1 is only 10-15% higher than the good cell break-even α_1 . This can be interpreted as it being rational to choose to build a system with bad cells if their marginal cost is less than roughly 85% that of the good cells.

This interpretation can be generalised: If the system designer has estimates of the degradation characteristics of good and bad cells, and of the costs associated with modularisation, and a revenue model, it is natural to ask how much less should they pay for the bad cells in order to justify choosing them over the good cells. Figure 4 aids a decision making process: it depicts the profit differential between an optimally designed system using bad cells and an optimally designed system using good cells (optimal in the sense of Figure 3) as the relative cost of the good versus bad cells and the modularisation cost k is varied. If the cost is positive, the optimal bad-cell system outperforms the optimal good-cell system and it is rational for the designer to choose to use the bad cell (and, by implication, smaller module sizes in order to handle the increased cell variability). An important result is that, even for high cost of modularisation ($k \rightarrow 10$) over long timeframes ($T \rightarrow 2$), it is necessary for bad cells to cost only about 22% less than the good cells to make their choice favourable.

5.4. Minimum capacity profit model. The second model is one where the system is required to offer at least some minimum capacity over a given lifetime, for example in an electric vehicle application where the manufacturer guarantees a certain number of miles-per-charge during a warranty period. In this case, the ACF plots of Figure 2 may be used to design a suitable system. To find the actual system capacity at any instant in time, the mean capacity of the population of cells (subfigure A) should be multiplied by the ACF at that instant. For example at time $T = 1$ the mean capacity of the cells is approximately 0.8. Therefore, for good cells, to deliver a system $C/n > 0.6$ at $T = 1$ requires an $ACF > 0.6/0.8 = 0.75$ and so only systems with $\ell \leq 625$ are viable.

In many applications, the penalty associated with failing to meet a particular capacity-lifetime point is not absolute. Instead, the manufacturer may pay a fixed penalty when the specification is not met because they must then provide compensation, repair or replacement. For example, this model is broadly applicable to an EV manufacturer who warrants their vehicle's battery for a certain time period and/or a maximum mileage. In this case the manufacturer should rationally trade the expected cost incurred by honouring the warranty against the up-front cost of installing larger battery packs (or choosing smaller ℓ) to ensure more headroom for cell degradation. This can be modelled as follows. First, define a warranty penalty

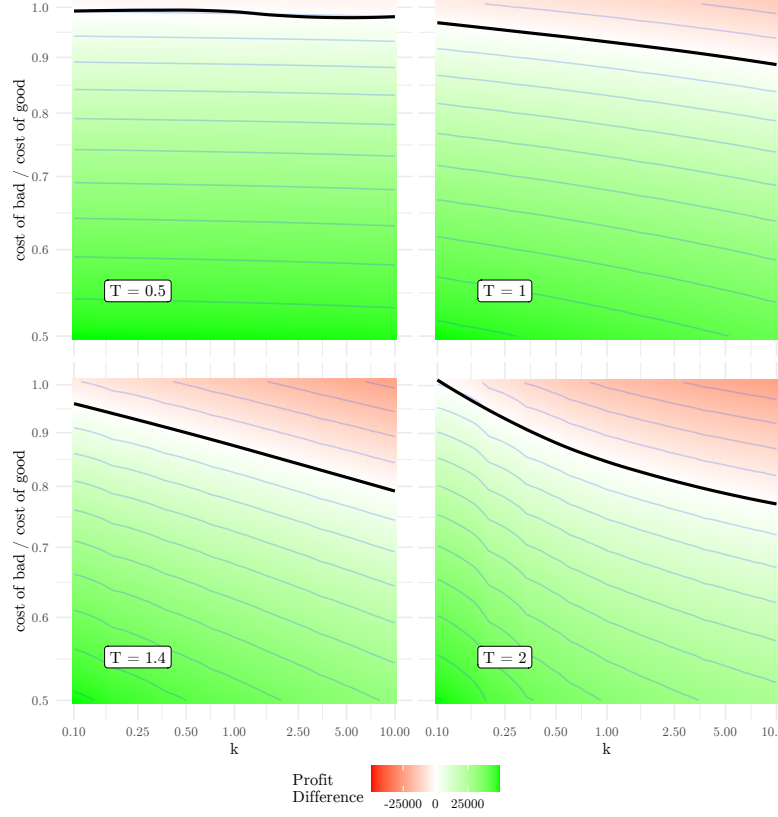


FIGURE 4. The difference in profit between a system using bad cells versus good as a function of the relative cost of bad to good cells and the relative cost of power electronics to cells. The thick black line represents break-even.

cost:

$$(12) \quad W(\ell, m) := \alpha_2(\ell, m) I_{C(T; \ell, m) < c^*}$$

where $I_{C(T; \ell, m) < c^*}$ is the indicator function of the event $C(T; \ell, m) < c^*$, i.e. it is 1 when $C(T; \ell, m) < c^*$ occurs, and 0 otherwise. $\alpha_2(\ell, m)$ is the penalty cost when warranty specifications are not met. We can then define the expected lifetime profit as follows (note that all terms are negative as we only consider costs in this model):

$$(13) \quad P_2(\ell, m) := \mathbb{E}[-K(\ell, m) - W(\ell, m)]$$

$$(14) \quad = -K(\ell, m) - \alpha_2(\ell, m) P(C(T; \ell, m) < c^*)$$

The up-front cost $K(\ell, m)$ of the system was defined earlier in Equation (8).

This helps illustrate a secondary value of modularisation beyond simple improvements in ACF: better predictability of system capacity over time despite uncertain cell degradation behaviour. The 95% confidence intervals of Figure 2 suggest that if $\ell > 2500$ the spread in system capacities becomes negligible, i.e. the negative systemic effects of cell capacity variation cease to be a concern because the system can be relied upon to deliver a tightly-defined ACF. In this case the battery system

can be designed assuming a simple mean-capacity model of the cells (of course, the underlying variation in cell degradation behaviour must be well-modelled in order to turn this into a prediction of actual system capacity and corresponding $W(\ell, m)$). The broad expectation is that successful system designs will give predictable capacities under a spread of cell degradation behaviour, i.e. will have an appropriately low ℓ for a given cell type.

In this context, the goal of the system designer is to minimise K subject to the constraint

$$(15) \quad \ell m = n$$

Once more, this is an integer programming problem in ℓ and m , although in this instance n is an optimisation variable (system capacity at a given time can be increased through better modularisation, but also by simply increasing the number of cells in the system). A reasonable choice is to set $\alpha_2 = K(\ell, m)$, representing the condition where failure to provide the specified minimum capacity over the specified lifetime results in a penalty equal to the cost of the battery system. In this case:

$$(16) \quad P_2(\ell, m) = -K(\ell, m)(1 + P(C(T; \ell, m) < c^*))$$

Figures 5 and 6 show the results of minimising 16, i.e. system designs in terms of the number of cells ℓm and module size ℓ that results in the lowest P_2 cost, for systems composed of either good or bad cells. Clearly, not only the minimum *cost*, but also the optimum *design* will depend on the cost of modularisation and the relative cost of good to bad cells. The bold boundary line running bottom-left to top-right allows the designer to make a simple binary decision: Given knowledge of modularisation cost and relative cell costs, is the lowest-cost system to be built using good or bad cells? The colouring of Figure 5 then indicates how many cells the lowest-cost system at any given design point contains. The results conform to expectations: if bad cells cost nearly as much as good cells, it is rational to select good cells as this will allow larger module sizes and so reduce modularisation costs. As the relative cost of bad cells goes down, a higher cost of modularisation can be tolerated whilst still leading to a lower total cost than for a system composed of good cells. This effect is quite strong: if bad cells are two-thirds the price of good cells, a bad-cell system is favourable until the modularisation cost k rises above 10, indicating that it is rational to invest in substantial amounts of additional electronics in order to manage cell variability (as opposed to investing in good cells to reduce variability directly).

For a fixed k (moving left to right on Figures 5 and 6), the optimal *number* of modules or cells in a design is not influenced by the ratio of good to bad cell cost except at the change point when bad cell systems become preferable. This is because the cost of modularisation is normalised to the cell cost; changing the cell cost scales the modularisation cost proportionally. The striations visible on the graph are a combined artefact of a discrete optimisation and the fact that a stochastic algorithm with limited runtime was to be used to locate the optimum design at each point (i.e. some minor variation in the solution from run-to-run should be expected).

The systems containing the least number of cells, and so the physically smallest, are clustered in the bottom of both Figures. As expected, these have small module sizes (i.e. many modules in Figure 6) in order to obtain high ACF throughout their lifetimes. There are six system designs that are the very ‘smallest’ in the sense that

they contain the minimum number of cells out of all the designs: these use 1294 good cells and are fully modularised (one cell per module), and are indicated with gold stars in Figure 5. In practice, the physical size of the module electronics would need to be considered in order to determine the true lowest-volume (or lowest-mass) system, but of course this is highly implementation-dependent.

Conversely, there are some designs described in Figures 5 and 6 that are likely to be infeasible in practice because they contain a lot of cells. Systems containing 3000 or more cells are highlighted with a dark shaded box in Figure 5. These systems occur in the top-right of the Figures where bad cells are very inexpensive but the cost of modularisation is high; in this case the lowest-cost design deploys many bad cells arranged in a small number of modules. In such a design, the degradation of a single cell will cause a low ACF but this can be tolerated simply because there are many cells all contributing a small amount of capacity. This implies a physically-large battery that would have major negative effects on the overall EV performance in terms of mass and acceleration (and would produce integration challenges), and so would not be viable in practice. The optimisation problem could be extended to place a penalty on total mass or volume in order to model such knock-on system effects.

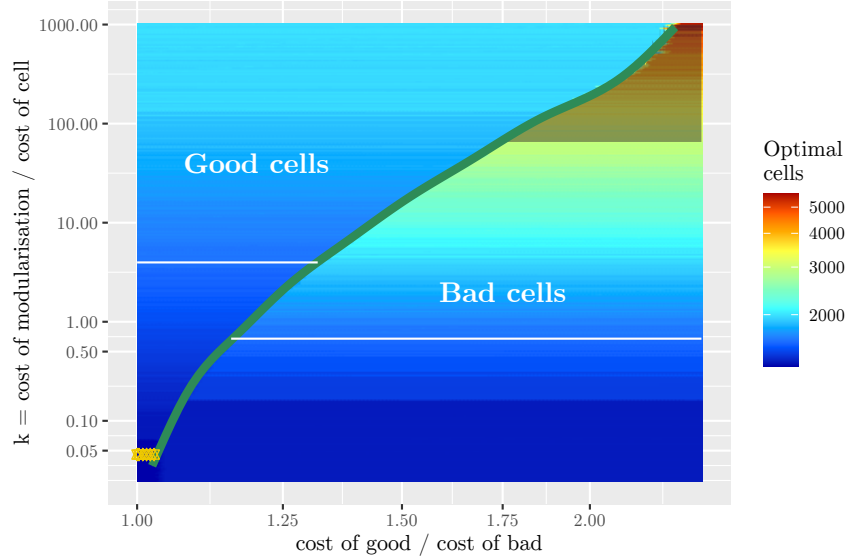


FIGURE 5. The number of cells in the lowest-cost system as defined by Equation (16), evaluated for a range of relative cell costs and relative modularisation costs. Each design is either composed entirely of good or bad cells, and the lowest cost design is chosen at each point. The boundary between both systems is indicated by the bold line. Systems below the white line contain 1600 cells or less.

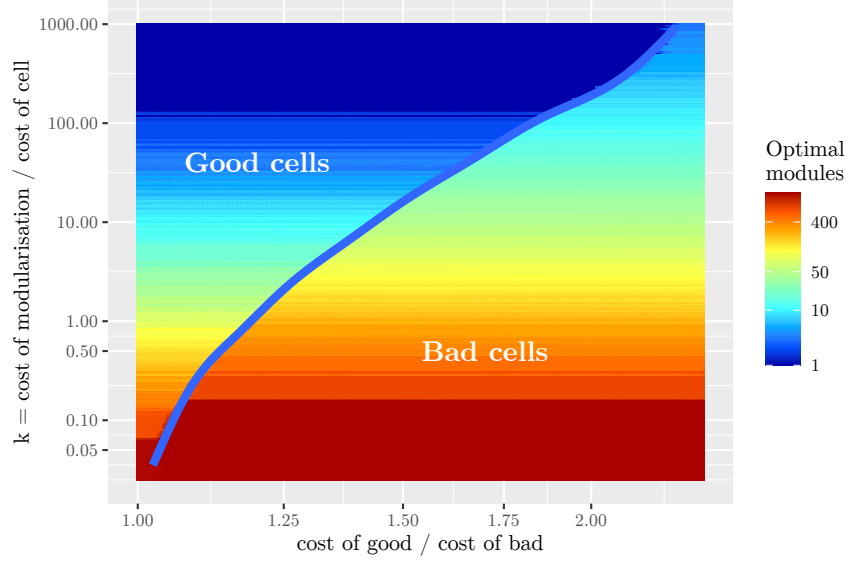


FIGURE 6. The number of modules in the lowest-cost system as defined by Equation (16), evaluated for a range of relative cell costs and relative modularisation costs.

6. CONCLUSION

This paper has explored the effect of cell variability and degradation on battery system design with a focus on the application of modularisation as a means to lessen the impact of cell variability on overall system performance. Underpinning the study is a simple statistical model designed to qualitatively exhibit similar behaviour to that observed in several large-scale cell cycling studies. Two measures of system performance were proposed, ACF and AICF, which provide an intuitive exposition of the advantages of using small module sizes.

In practice, however, a battery system designer is faced with an essentially economic decision: What is the choice of cell type and system design that meets a specification at minimum cost? We have shown how these two choices are inter-linked; the designer cannot choose the best system design without knowing how the cell population is likely to degrade, and they may decide to use a different cell type depending on non-cell costs (i.e. the costs of modularisation). This is essentially summarised by the observation that if supporting electronics are cheap, it may be rational to use cheaper-but-higher-variability cells arranged into many, small modules.

In an attempt to provide answers to such economics-driven decision problems, two example cost models were proposed and solved to find the lowest-cost system designs. The first model was a simple lifetime capacity maximisation applicable to scenarios in which the accumulated lifetime system capacity is a good indicator of system value, which we suggest is a fair model of grid-connected energy storage systems. The second model was intended to represent scenarios where some minimum level of capacity must be maintained over a given lifetime, which we suggest aligns well with an EV manufacturer providing a range warranty.

For the ACF and AICF measures, and the two economic models, the broad conclusion is that very significant cell variability (e.g. that of the bad cells in this paper) can be compensated to a great degree by suitably fine-grained modularisation (roughly at the 10-cell level). It would be instructive to further study the real costs associated with modularisation, and for battery system engineers to propose methods of achieving fine-grained modularisation in a cost-effective manner.

As discussed in Section 1, the results presented in this paper are broadly applicable to *any* ESS that can create different charge/discharge rates between defined groups of cells. This includes all systems that are truly modular (i.e. that can fully bypass cells as required and so have an infinite differential-rate capability), but also includes all systems with balancing circuitry, under the assumption that a given system's maximum differential-rate capability is not exceeded. This assumption translates directly into a limit on the maximum variation in cell capacities that a given system can tolerate. If this assumption holds, then the system will behave as predicted here. If it does not hold, then the ACF and AICF of the system will be limited by the particular differential-rate capability of that system for the time when the capability is exceeded.

REFERENCES

- [1] T. Baumhöfer, M. Brühl, S. Rothgang, and D. U. Sauer. Production caused variation in capacity aging trend and correlation to initial cell performance. *Journal of Power Sources*, 247:332–338, 2014. doi:10.1016/j.jpowsour.2013.08.108.
- [2] R. Billinton and R. N. Allan. *Reliability Evaluation of Engineering Systems: Concepts and Techniques*. Plenum Press, 2nd edition, 1992.
- [3] R. Billinton and R. N. Allan. *Reliability Evaluation of Power Systems*. Plenum Press, 2nd edition, 1996.
- [4] C. R. Birkel, M. R. Roberts, E. McTurk, P. G. Bruce, and D. A. Howey. Degradation diagnostics for lithium ion cells. *Journal of Power Sources*, 341:373–386, 2017. ISSN 0378-7753. doi:10.1016/j.jpowsour.2016.12.011.
- [5] E. Chatzinikolaou and D. J. Rogers. Cell soc balancing using a cascaded full-bridge multilevel converter in battery energy storage systems. *IEEE Transactions on Industrial Electronics*, 63(9):5394–5402, Sep. 2016. ISSN 0278-0046. doi:10.1109/TIE.2016.2565463.
- [6] E. Chatzinikolaou and D. J. Rogers. A comparison of grid-connected battery energy storage system designs. *IEEE Transactions on Power Electronics*, 32(9):6913–6923, Sep. 2017. ISSN 0885-8993. doi:10.1109/TPEL.2016.2629020.
- [7] E. Chatzinikolaou and D. J. Rogers. Hierarchical distributed balancing control for large-scale reconfigurable ac battery packs. *IEEE Transactions on Power Electronics*, 33(7):5592–5602, July 2018. ISSN 0885-8993. doi:10.1109/TPEL.2017.2782675.
- [8] M. H. DeGroot. *Optimal Statistical Decisions*. Wiley, 2004. doi:10.1002/0471729000.
- [9] A. Devie, G. Baure, and M. Dubarry. Intrinsic variability in the degradation of a batch of commercial 18650 lithium-ion cells. *Energies*, 11:1031, 04 2018. doi:10.3390/en11051031.

- [10] W. Diao, N. Xue, V. Bhattacharjee, J. Jiang, O. Karabasoglu, and M. Pecht. Active battery cell equalization based on residual available energy maximization. *Applied Energy*, 210:690 – 698, 2018. ISSN 0306-2619. doi:<https://doi.org/10.1016/j.apenergy.2017.07.137>. URL <http://www.sciencedirect.com/science/article/pii/S0306261917310115>.
- [11] A. Farmann, W. Waag, and D. U. Sauer. Application-specific electrical characterization of high power batteries with lithium titanate anodes for electric vehicles. *Energy*, 112:294–306, 2016. doi:10.1016/j.energy.2016.06.088.
- [12] A. Friesen, F. Horsthemke, X. Mönnighoff, G. Brunklaus, R. Krafft, M. Börner, T. Risthaus, M. Winter, and F. M. Schappacher. Impact of cycling at low temperatures on the safety behavior of 18650-type lithium ion cells: Combined study of mechanical and thermal abuse testing accompanied by post-mortem analysis. *Journal of Power Sources*, 334:1–11, 2016. doi:10.1016/j.jpowsour.2016.09.120.
- [13] D. F. Frost and D. A. Howey. Completely decentralized active balancing battery management system. *IEEE Transactions on Power Electronics*, 33(1): 729–738, Jan 2018. ISSN 0885-8993. doi:10.1109/TPEL.2017.2664922.
- [14] X. Han, M. Ouyang, L. Lu, and L. Jianqiu. Cycle life of commercial lithium-ion batteries with lithium titanium oxide anodes in electric vehicles. *Energies*, 7:4895–4909, 08 2014. doi:10.3390/en7084895.
- [15] S. J. Harris, D. J. Harris, and C. Li. Failure statistics for commercial lithium ion batteries: A study of 24 pouch cells. *Journal of Power Sources*, 342:589–597, 2017. doi:10.1016/j.jpowsour.2016.12.083.
- [16] S. Lehner, T. Baumhöfer, and D. U. Sauer. Disparity in initial and lifetime parameters of lithium-ion cells. *IET Electrical Systems in Transportation*, 6 (1):34–40, 2016. ISSN 2042-9738. doi:10.1049/iet-est.2015.0007.
- [17] D. V. Lindley. *Making Decisions*. Wiley, London, 2nd edition, 1985.
- [18] M. Liu, W. Li, C. Wang, M. P. Polis, L. Y. Wang, and J. Li. Reliability evaluation of large scale battery energy storage systems. *IEEE Transactions on Smart Grid*, 8(6):2733–2743, Nov 2017. ISSN 1949-3053. doi:10.1109/TSG.2016.2536688.
- [19] Z. Liu, C. Tan, and F. Leng. A reliability-based design concept for lithium-ion battery pack in electric vehicles. *Reliability Engineering and System Safety*, 134:169–177, 2015. ISSN 0951-8320. doi:10.1016/j.ress.2014.10.010.
- [20] L. Maharjan, T. Yamagishi, and H. Akagi. Active-power control of individual converter cells for a battery energy storage system based on a multilevel cascade pwm converter. *IEEE Transactions on Power Electronics*, 27(3):1099–1107, March 2012. ISSN 0885-8993. doi:10.1109/TPEL.2010.2059045.
- [21] T. Morstyn, M. Momayyezani, B. Hredzak, and V. G. Agelidis. Distributed control for state-of-charge balancing between the modules of a reconfigurable battery energy storage system. *IEEE Transactions on Power Electronics*, 31 (11):7986–7995, Nov 2016. ISSN 0885-8993. doi:10.1109/TPEL.2015.2513777.
- [22] C. P. Robert and G. Casella. *Introducing Monte Carlo Methods with R*. Springer, 2010.
- [23] K. Rumpf, M. Naumann, and A. Jossen. Experimental investigation of parametric cell-to-cell variation and correlation based on 1100 commercial lithium-ion cells. *Journal of Energy Storage*, 14:224–243, 2017. ISSN 2352-152X. doi:10.1016/j.est.2017.09.010.

- [24] K. A. Severson, P. M. Attia, N. Jin, N. Perkins, B. Jiang, Z. Yang, M. H. Chen, M. Aykol, P. K. Herring, D. Fraggadakis, M. Z. Bazant, S. J. Harris, W. C. Chueh, and R. D. Braatz. Data-driven prediction of battery cycle life before capacity degradation. *Nature Energy*, 4(5):383–391, May 2019. ISSN 2058-7546. doi:10.1038/s41560-019-0356-8.
- [25] B. Stiaszny, J. C. Ziegler, E. E. Krauß, J. P. Schmidt, and E. Ivers-Tiffée. Electrochemical characterization and post-mortem analysis of aged $\text{limn2o4-li(ni0.5mn0.3co0.2)o2/graphite}$ lithium ion batteries. part i: Cycle aging. *Journal of Power Sources*, 251:439–450, 2014. ISSN 0378-7753. doi:10.1016/j.jpowsour.2013.11.080.

APPENDIX A. OPTIMAL ORDERING OF CELLS

Theorem 1. *Consider any non-decreasing sequence of real numbers c_0, \dots, c_{n-1} , where $n = m \times \ell$. Let \mathcal{S} be the set of all possible permutations of $\{0, \dots, n-1\}$. Then*

$$(17) \quad \sum_{i=0}^{m-1} \ell \min_{j=0}^{\ell-1} c_{\ell i+j} = \max_{\sigma \in \mathcal{S}} \sum_{i=0}^{m-1} \ell \min_{j=0}^{\ell-1} c_{\sigma(\ell i+j)}.$$

Proof. Consider any permutation σ . We can bring $c_{\sigma(0)}, \dots, c_{\sigma(n-1)}$ into non-decreasing order through selection sort. In selection sort, we swap the first element of the sequence with the sequence’s minimal value. Then, we swap the next element with the minimal value of the sequence excluding the first value, and so on. We show that every step of the selection sort algorithm applied on the sequence cannot decrease the value. Because we can do this for every permutation σ , it follows that the value is maximal when the sequence is non-decreasing.

To make the proof concise, we do not apply selection sort to the letter, as in each step we will perform some additional swaps in the yet to be sorted part of the sequence. This will still result in a fully ordered sequence at the end of the algorithm: all we need for the proof to work is that the algorithm builds up a fully sorted sequence.

For brevity of notation, let $d_p := c_{\sigma(p)}$ for all $p \in \{0, \dots, n-1\}$. The value of the objective function is then

$$(18) \quad \sum_{i=0}^{m-1} \ell \min_{j=0}^{\ell-1} d_{\ell i+j}.$$

Suppose that the first r values of the sequence d_0, \dots, d_{n-1} have been sorted already through selection sort, i.e.

$$(19) \quad d_0 \leq d_1 \leq \dots \leq d_{r-1} \leq \min_{p=r}^{n-1} d_p$$

Without loss of generality, we can assume that for every $i \in \{0, \dots, m-1\}$

$$(20) \quad d_{\ell i} = \min_{j=0}^{\ell-1} d_{\ell i+j},$$

and that

$$(21) \quad d_0 \leq d_{\ell} \leq d_{2\ell} \leq \dots \leq d_{(m-1)\ell}$$

because we can always achieve this with some simple swaps which leave the value of the objective function unchanged, and which also leave the first $r - 1$ values unchanged.

Selection sort proceeds by swapping d_q and $d_{q'}$, where q' is such that $d_{q'} = \min_{p=q}^{n-1} d_p$. Let $i, i' < m$ and $j, j' < \ell$ be the unique numbers such that $q = \ell i + j$ and $q' = \ell i' + j'$.

If $j = 0$, then we can choose $q' = q$, that is, we do not need to swap anything. Indeed, because of the assumptions we made in Equations (20) and (21), $d_{\ell i}$ is already equal to $\min_{p=\ell i}^{n-1} d_p$.

If $i = i'$, then the swap will also clearly leave the objective function unchanged.

If $j > 0$, and $i < i'$, then by construction $d_{\ell i} \leq d_{\ell i'} \leq d_{\ell i' + j'}$, and so

$$(22) \quad \ell \min_{q=0}^{\ell-1} d_{\ell i + q} = \ell d_{\ell i} = \ell \min \left\{ \min_{q=0, q \neq j}^{\ell-1} d_{\ell i + q}, d_{\ell i' + j'} \right\}.$$

We also have that, $d_{\ell i' + j'} \leq d_{\ell i + j}$, and therefore

$$(23) \quad \ell \min_{q=0}^{\ell-1} d_{\ell i' + q} \leq \ell \min \left\{ \min_{q=0, q \neq j'}^{\ell-1} d_{\ell i' + q}, d_{\ell i + j} \right\}.$$

We conclude that the value of the objective function cannot decrease. \square