The award of the UK digital dividend cleared spectrum

Issues in auction design

A report for Ofcom

31 May 2008

Executive Summary

1. This paper considers auction design approaches for Ofcom's forthcoming award of cleared digital dividend spectrum. Allocation of this spectrum by way of an auction is particularly challenging owing to the structure of the available spectrum and the need to create a level playing field for a number of different potential uses that could co-exist within the band.

2. Previous work by Ofcom, supported by DotEcon and other consultants, has already established that an auction is appropriate for the cleared spectrum. There are a wide range of possible uses for this spectrum and attendant high risk of regulatory failure associated with pre-judged outcomes. Accordingly, it is essential that the auction format allow the widest range of uses and users to compete in an auction on a reasonably level playing field. This is consistent with Ofcom's objective for the DDR, which is "*to maximise the total value to society that using the digital dividend generates over time*."[1]

3. As with all other Ofcom awards, it is emphatically not Ofcom's objective to award the digital dividend so as to maximise revenue for the Exchequer. Given Ofcom's duties, this is not a consideration that we take into account in making our recommendation on the optimal auction format. Our focus is on achieving efficient outcomes.

4. Our main findings and recommendations are as follows:

   a) Ofcom should adopt a combinatorial clock auction (CCA) format for this award;

   b) Lots should be divided up into multiple categories, tailored to each of the five possible usage groups that might potentially use this spectrum;

   c) The auction rules (such as activity rules and rules for increasing prices during the auction) must be tailored to the specific lot structure chosen;

   d) Ofcom's proposed lot structure, as discussed in this paper and outlined in detail in Ofcom's latest consultation, is feasible but challenging to implement in terms of the computing resources needed to find the optimal use of the spectrum;

   e) If it were possible to simplify the lot structure without creating undue uncertainty for bidders, this would be advantageous as it would significantly ease implementation of the auction.

   These findings are summarised below and explained in detail in this paper. In addition, DotEcon has prepared advice on draft auction rules, which has been reflected in Ofcom's latest consultation.

The CCA format

5. We recommend that Ofcom adopt a combinatorial clock auction (CCA) format for this award. This would be similar in basic structure to the auction

[1] Ofcom, Statement on the Digital Dividend Review, December 2007.

formats used/proposed for the UK 10-40GHz, L-band and 2.6GHz auctions, and Dutch 2.6GHz auction.

6. In reaching this recommendation, we have considered a wide range of other auction formats based on the simultaneous multiple round auction (SMRA) family of auction formats. Possible alternatives included the 'standard' SMRA format (widely used for 3G awards and North American spectrum auctions), the SMRA with augmented switching (used for the 2.6GHz awards in Norway and Sweden) and the SMRA with package bidding (proposed for use by the US FCC but never implemented).

7. We believe that the CCA is much more likely to produce an efficient auction outcome than any of the other auction formats. There are a number of reasons for this:

- A key issue in this case is the need to ensure appropriate separations between different potential uses on an efficient basis. Only a package bidding auction, such as the CCA, can guarantee such separations. With more standard SMRA formats, bidders cannot be certain they will win usable spectrum, given that they do not know who their neighbour will be.
- Package bidding is also very important in this auction in order to alleviate aggregation risks for bidders; for example, guaranteeing contiguity of lots where appropriate and allowing bidders to manage demand for substitute and complementary lots across the band. These aggregation risks vary for different types of use and user; therefore, it is not feasible to address them through packaging. More standard SMRA formats expose bidders to winning fragmented spectrum or unwanted subsets of their demand.
- Although package bidding can introduce threshold risks for smaller bidders, we do not think this a major concern for this auction. The proposed second price rule provided reasonable incentives for all bidders to bid on the basis of their true values, especially where information about bids of others is not available. Further, given the diverse range of possible auction outcomes, in terms of the allocation of lots to different uses and users, the likelihood that bidders would be able to identify opportunities for strategic bidding behaviour that offer any significant likelihood of success, is small. Any such risks can be further reduced by only releasing information about the level of demand for lot categories, not information about individual bids.
- Although, in principle, the 'SMRA with package bidding' can also address aggregation risks and separation requirements in the same way as a CCA, our judgement is that this format is not practical in this context. Specifically, the winner and price determination algorithms are too complex to run on a round-by-round basis, and the package bid input requirements for bidders would be unduly onerous.

The lot structure

8. The main challenge with using a CCA for this award is created by the complexity of the packaging proposals that Ofcom have developed. There are many more types of lot categories and possible juxtapositions of dissimilar usage rights than occurred in any of the previous CCAs. The determination of whether any particular demand for lots in different categories can be fitted into the available spectrum is therefore a complex process. This requires a much more extensive set of rules to define what patterns of usage are feasible than has been needed in previous awards.

9. The complex lot structure means that there are a large number of possible ways in which the available spectrum could be allocated to the different candidate lot categories that have been defined by Ofcom (i.e. different 'band plans'). It follows that the detailed auction rules can only be finalised when the lot categories have been finalised. Although we have evaluated the implications for the auction of the lot categories described in Ofcom's consultation document, we understand that this categorisation may be subject to change as a result of the consultation. Accordingly the corresponding auction rules may need to be revised in light of the consultation and further analysis.

10. In this respect, the introduction of additional lot categories should in our view be resisted as it would further increase complexity and could make the auction very difficult or, ultimately, even infeasible to run. Conversely, it would highly beneficial to collapse some of the potential categories of lots we have identified if it were found that the effect of these distinctions on bidders' valuations are likely to be modest.

Auction rules

11. Because we recommend that Ofcom adopt an auction structure that is fundamentally similarly to that used for the 10-40GHz and 2.6GHz, with both a Principal Stage and Assignment Stage, many of the auction rules developed for these awards can be carried over to the rules for this award. Nevertheless, there are a series of specific issues with the auction rules for the DDR cleared award that will require further analysis or decisions. The most important ones that we have identified are:

- The round-by-round pricing rule, which is necessarily more complex that previous CCAs owing to the potential for overlapping demand for the same spectrum from bidders requiring different lot categories. We have made outline proposals for calculating excess demand for the purpose of administering the auction. However these need further development and validation to ensure that the general approach can be customised to work well in this particular case when the packaging details are finalised.
- The activity rule – while we have so far explored with Ofcom how a simple eligibility points rule could be applied in the context of the DDR cleared award, there may be efficiency benefits from switching to an alternative rule based on revealed preference give the particular

characteristics of this award (especially that there are many categories of lot with uncertain relative value). Further work would be needed to develop and validate any such rule.

- The procedures for offering specific frequencies to successful bidders in the Principal Stage. Depending on the approach taken, this could result in a much more complex assignment bidding stage than has been developed for other CCAs (such as the 10-40 GHz award). It may be possible to impose auction rules to limit the possible combinations that need to be considered in the assignment phase without significant detriment to the efficiency of the outcome. However, the proper specification of such rules requires further detailed work, based on the finalised specification of lot categories. These could be significantly affected by the desired level of flexibility applied to rules governing FDD duplex separations (e.g. a move to fixed-frequency rasters).

Feasibility assessment

12. Our conclusion is that it is feasible to use the lot structure described in Ofcom's consultation document in a CCA auction for the DDR cleared award, and to solve explicitly and unambiguously for the winning bidders. However, this problem is challenging in computing terms given the large numbers of possible outcomes in terms of how spectrum might be used, which is turn is creates by the large number of lot categories proposed. This is the downside of taking a flexible, market-driven approach.

13. Without the use of high performance computing techniques, it is not be feasible to guarantee solving this problem within a reasonable time. This will affect the costs of the necessary equipment, although we estimate that the problem can still be solved on a timely basis using a cluster of high performance commodity computers. Significant software development will also be necessary to develop algorithms currently used for winner determination for a parallel processing environment. On this basis we recommend that further development work be undertaken prior to finalisation of the lot categories and associated auction rules.

Annex 1: Assessing complexity

14. We consider the question of whether it is practical to run a CCA similar to that used or proposed by Ofcom for other spectrum bands. With a large amount of spectrum and many potential competing uses, the auction will be complex. There are two aspects to this complexity:

- how complicated it may be for bidders to express their preferences through package bids; and
- how difficult it might be to evaluate the winning bids and winning prices.

15. The key concern here is the latter: the impact of the large numbers of usage right groups and categories on the complexity of the winner determination process. However, we start by consider the somewhat easier question of how difficult the auction might be for bidders.

Complexity for bidders

16. For bidders, even with the proposed finely differentiated lot structure, the auction should not be excessively complex. A typical bidder will not be interested in all the different groups of usage right (e.g. one bidder is unlikely to be interested in both DVB-T and FDD uses). Therefore, a typical bidder will only be interested in a tiny fraction of the packages potentially available.

17. For a bidder, the complexity of the auction is related to the number of categories of lots that the bidder considers are close substitutes. This number determines the need to prepare multiple package bids running through the various possible ways in which the bidder's requirements could be met from either one category or another. The number of categories for each usage right group is shown in Table 1.

Table 1: Number of categories by usage right group

Usage right group	Categories in lower sub-band	Categories in upper sub-band	Total categories
MMS	6	4	10
DVB-T	6	3	9
TDD	2	1	3
FDD-uplink	2	1	3
FDD-downlink	5	5	10
Channel 38	1	-	1
TOTAL	**22**	**12**	**36**

18. The usage right groups that are most finely differentiated by frequency are the DVB-T and MMS usage right groups. In these cases, there are 6 categories in the lower sub-band and a further 3 or 4 categories in the upper sub-band (9 or 10 categories across both bands). This is fewer categories than in the L-band auction (where there were 17 categories, of which 16 were partial substitutes). Moreover, not all of these categories are likely to be substitutes for typical bidders; for example, categories in different DVB-T aerial groups are likely to be complements, rather than substitutes.

19. Some specific modelling will be needed to investigate how many distinct packages bidders with various different types of preferences might reasonably need to express those preferences. This is important as it will be necessary to cap the number of distinct packages on which bids can be entered for practical reasons (in particular, the number of possible packages is astronomical, so a limit needs to be imposed for the reliable functioning of the auction software). However, provided that most bidders do not wish to bid for packages of lots drawn from many different usage right groups, the number of packages likely to need bids should be significantly fewer than the L-band auction.

20. To keep the auction as simple as possible for bidders, it is important that additional categories are not introduced for lots that are close substitutes for most bidders. For example, suppose that a bidder is interested in four lots in one category. If this category is split into two categories, which the bidder finds largely identical, then it will be necessary for that bidder to make package bids across the various splits that might occur between the two categories (i.e. 0-4, 1-3, 2-2, 3-1 and 4-0). Therefore, if subsequent evidence suggests that some of the categories proposed might be collapsed, this would be useful in terms of simplifying the decisions that bidders need to make as well as the complexity of the winner determination problem.

Complexity in winner determination and pricing

21. Our main concern is with complexity on the auctioneer side and, in particular, the computational burden of determining who the winners are and what they should pay. Mathematically, the problem of finding prices is largely a matter of re-running the winner determination problem a number of times with modified notional bids.[2] Therefore, the practicality of both the winner determination and pricing steps reduces to the question of how long the winner determination step takes to achieve.

22. Although winner determination is a question about what feasible combination of bids is of greatest total value, this problem is not best solved by forming all the combinations of bids. The number of combinations of bids rapidly becomes very large as the number of bidders and the number of bids increases. This ‘blow up’ is exponential. For example, there are about a million combinations of 20 bids and about 1 billion combinations of 30 bids.

[2] See http://www.dotecon.com/publications/dp0701.pdf

We may need to solve a winner determination problem with hundreds or even thousands of bids; this has an astronomical number of bid combinations. Therefore, a more refined approach is needed.

23. By using more efficient algorithms[3] it is possible to solve the winner determination in a way where computer run times increase linearly with the number of bids and bidders (i.e. twice as many bidders takes roughly twice as long). However, these more efficient methods need to search across all the feasible patterns of usage that can occur within a band. Therefore, the determinant of computational complexity is the number of different number ways in which the band could be allocated to different usage rights.[4]

24. To avoid subsequent confusion, it is useful to define the following terms:

- An ***outcome*** is a description of who wins each block and what use each winner is permitted to make of that block;
- A ***usage pattern*** is a description of what each lot in the band will be used for, i.e. block by block which usage right category will apply to that block, including the possibility of a block being required as a guard block or being unallocated;
- A ***usage vector*** is list of numbers of lots of each usage right group and category.

25. Therefore, an outcome contains the most information. A usage pattern is simply a description of how the blocks will be used, but without any information about who gets what – in effect a band plan without actual users marked in. A ***feasible usage pattern*** is one that satisfies all of the constraints proposed about what usage rights are allowed to go where and with appropriate separations created by guard blocks between dissimilar usage rights. A ***feasible usage vector*** describes a certain ***number*** of lots from each usage right group and category than are able to fit together within the band; this does not include information about where the lots might be located at specific frequencies.

26. For certain numbers of lots from the different usage right categories to be a *feasible usage vector*, there has to some *feasible usage pattern* with the corresponding number of lots of each category. In this case, these lots can be packed together within the band in some way. Often, there will be many ways in which the lots can be packed into the band. Therefore, there will typically be many feasible usage patterns corresponding to one feasible usage vector.

[3] Ibid.

[4] If the number of bids is small, there are more efficient algorithms. However, we interested in the computation demands when the number of bids is large and an exact optimum must be found. In this case, the number of feasible usage vectors is the key determinant of run times. More efficient algorithms (in particular, algorithms that are not NP-hard) are available for finding approximate optima.

27. These concepts are closely related to the different functions of the principal stage and the assignment stage. To solve the winner determination problem in the principal stage, we maximise the sum of winning bids, subject to accepting at most one bid from each bidder and such that the total number of lots allocated across all the winners is a feasible usage vector. Therefore, to solve the principal stage winner determination problem, we do not need to know anything about how the usage rights might be located at particular frequencies, only that they can be packed into the band in some way or another. Therefore, to solve the winner determination problem, you need to know all the feasible usage vectors, but you do not need to know all the feasible usage patterns.

28. In fact, it is the number of feasible usage vectors that determines the overall computational load. This is shown in the in the Annex. Using efficient methods, both the run times and the peak memory requirements are proportional to the number of feasible usage vectors (at least in the case of many bids and bidders)[5]. Therefore, to assess the feasibility of solving the winner determination problem, we need to find the number of feasible usage vectors corresponding to the lot structure. This is typically a much smaller number (by orders of magnitude) that the number of distinct usage patterns or 'band plans'.

29. In general, increasing the number of lot types produces an exponential blow-up in the number of usage vectors. For example, suppose that there are ten blocks. If these are treated as identical, there are only eleven feasible usage vectors (i.e. between zero and ten lots in this single category). If we divide these into two categories with five lots each, there are 36 feasible usage vectors. If each lot is treated as a distinct category in its own right (the most extreme case) there are 1,024 feasible usage vectors. Therefore, to avoid complexity we should avoid making distinctions between different categories of usage right by location within the band whenever we can.

Computational hardware requirements

30. To give some context, the L-band auction is the most complex of the other CCAs currently planned or undertaken by Ofcom in terms of working out the winners. There are roughly a million feasible usage vectors in this case. With typical examples, solving for the winning bids and the prices to be paid takes a few minutes. With many thousands of bids, it might take an hour or more using an ordinary PC. In practice, the outcome of the actual auction was trivial to solve. However, clearly we must be able to handle the worst case.

31. As the number of feasible usage vectors increases, the primary difficulty is that intermediate working results (so-called value functions) need to be held

[5] Although in practice it is possible there might not be many bids, the proposed methods need to work robustly in the case of many bidders and many bids.

in computer memory. In a typical implementation, if there are *N* feasible usage vectors, then it is necessary to use 4*N* bytes of memory to hold intermediate results (assuming 64-bit arithmetic). It is quite possible to use computers with multiple processor cores or to use clusters of computers to speed up the calculations as the algorithms are very convenient for parallel processing methods. However, the way the calculations are structured, it is necessary to have one working table of size 4*N* held in memory and simultaneously accessible for all the processors to read and perform calculations on to produce a new table of size 4*N*. Therefore, provided that it is possible to hold a table of size 4*N* in a way that it can be efficiently accessed by many processors, current techniques can be scaled up to use multi-processor architectures. Once 4*N* becomes too large, this is not possible and there are then considerable overheads to moving data back and fore across multiple processors. This does not mean that the problem is impossible to solve at that scale, rather that it is very difficult without specialist hardware and a compete reworking of the algorithms currently used.

32. A very rough rule of thumb for the implications of various orders of magnitude of the number of feasible usage vectors in given in Table 4 below. This is based on our experience of how longer simpler problems might typically take to solve. Problems with tens or hundreds of millions of feasible usage vectors can be approached by scaling up the algorithms used for L-band to run across multiple processor cores. This means a significant expense in equipment, as it is already simply infeasible to solve such problems without clustering many computers.

33. Once the number of feasible usage vectors exceeds a few billion, we hit a fundamental limitation. It is no longer possible to keep using the same algorithm and control run times by using more processor cores. Commodity hardware with a 64-bit operating system can address at most 32GB of RAM. Therefore, once we hit tens of billions of feasible usage vectors, it is impossible for a commodity computer to hold an entire working table in RAM. At this point, the problem becomes a major challenge requiring supercomputing hardware (i.e. processors with fast interconnects) due the amount of data than needs to be passed back and for between processors. However, even before this point, the hardware costs of solving this problem in a reasonable time are very considerable.

Table 2: How tough is it to solve for winners?

Number of feasible usage vectors	Hardware requirements for typical worst-case scale of problem encountered in an auction
$<10^5$ (10-40GHz, 2.6GHz auctions)	Results within a few seconds to a few minutes on a typical PC
10^6 (L-band)	Results with a few minutes to an hour on a typical PC
10^7	Fast multiprocessor machine required Results within a few hours
10^8	Fast multiprocessor (8-way) machine required Results within a day
10^9	Cluster of multiprocessor machines needed 64-bit operating system now essential due to memory requirements
10^{10}	Large cluster of multiprocessor machines needed Results may take days to compute Limit of working tables being held fully accessible in memory to a single processor on commodity machinery
10^{11}	Specialist supercomputing hardware required as working tables held in distributed storage

Calculating the amount of complexity

34. How many feasible usage vectors are there in the auction based on the current Ofcom proposals? To take account fully of all the various constraints on how spectrum might be used would be a time-consuming exercise requiring software to be built to answer this question. Therefore, we have simplified the problem and calculated an upper bound on the number of feasible usage vectors by only taking account of the most important constraints. A more detailed description of the mathematics is provided in the Annex.

35. A key feature of this auction is that dissimilar uses require separations, which, in some cases, are considerable. This limits the number of outcomes with lots in many different usage right groups being awarded simultaneously. This is a very helpful feature that limits the overall complexity and forms is the heart of the method we deploy for estimating the number of feasible usage vectors.

36. The starting point is to notice that the feasible usage patterns for the lower and upper sub-bands are unrelated. How one sub-band is used does not affect the constraints on how the other might be used. Given this, we start

by considering the upper sub-band first as this is somewhat simpler (the lower sub-band is complicated by the presence of channel 38).

37. We can cap the number of feasible usage vectors through the following procedure:

1. Generate all the different cases of presence/absence of each usage right group in the band. There are potentially 32 of these (i.e. 2^5), but as FDD-uplink cannot be present without FDD-downlink, this reduces the number of cases to 24;
2. In each of these cases, we can calculate the most efficient way of ordering the usage rights in the band in order to minimise the total number of guard blocks between dissimilar uses.[6] Assume that similar uses are kept together in contiguous frequency blocks. Then imagine swapping around the order of the different usage rights within the band. This potentially changes the separations required. In addition, certain usage right groups (TDD and FDD-uplink) sometime require separations from out-of-band uses and so cannot be placed too close to an end. We can find the total separation required and organise the order of the five usage rights to minimise this. Notice that this minimum total separation is a function of which usage rights are present in the band, not how much spectrum is allocated to each use. In practice, there may be other constraints that we have not taken into account that may increase the number of guard blocks required beyond this minimum;
3. Given a case of the presence/absence of each usage right, we know the maximum amount of spectrum potentially available for allocation. This is the total amount of spectrum in the sub-band, less the minimum total separation required with this mix of uses. This is the maximum amount of spectrum that could possibly be allocated to the uses that are present;
4. We can find all the various ways feasible of splitting the spectrum available for allocation into five usage right groups, plus an additional category for unallocated spectrum. This gives a list of all the different combinations of total numbers of lots within each usage right group that are consistent with awarding no more than the maximum possible amount of available spectrum given this mix of uses.
5. For each possible combination of total numbers of lots within each usage right group, look at the number of ways of breaking the lots in a usage right group across the available categories of lot within that group.

38. Notice that this method does not take account of all possible constraints on how lots need to pack into the available spectrum and so provides an upper bound on the number of feasible usage vectors. In particular, we have assumed that in order to minimise the amount of spectrum in guard blocks, all usage rights can be moved around.

[6] We do not need to consider TDD-TDD separation here, as we are only concerned with which use particular blocks have, not who wins them.

39. Working out whether or not a particular number of lots of certain types will or will not pack into the available spectrum is computational quite tough.[7] It would be very time consuming to screen all the possible usage vectors generated by the procedure above to see whether or not each would pack into the available spectrum. Therefore, we take the easier step of simply checking whether allocating a certain number of lots in each category would allocated too many blocks either in the sub-band as a whole or in any particular frequency sub-range within the sub-band. The procedure is detailed in the Annex. The conditions we impose are necessary but not sufficient for a usage vector to be feasible, so we still obtain an overestimate of the number of feasible usage vectors. Nevertheless, the number is sufficiently reduced that it will be possible in a future exercise to obtain the exact list of feasible usage vectors by starting from this much-reduced list and checking each one for whether or not it can be accommodated in the available spectrum.

40. The procedure for the lower sub-band is somewhat similar. However, matters are somewhat complicated by the presence of a low-power usage right at channel 38. We can deal with this by first thinking about how channel 39 and 40 can be allocated. There are a limited number of possibilities, as there is insufficient spectrum above channel 38 for dissimilar uses to co-exist:

- All spectrum above channel 38 could be unallocated;
- There could be FDD-downlink lots, of which there are three different categories with a single lot available. Therefore, there are seven cases corresponding three different categories of single lots each either being allocated or not allocated (excluding the here the case of all spectrum above channel 38) being unallocated;
- There could be DVB-T lots, of which there are two single lot categories. Therefore, there are three possibilities according to which are allocated or not;
- There could be MMS lots, of which of which there are two single lot categories. Again, there are three possibilities according to which are allocated or not.

This gives 14 cases in total for how the spectrum in channels 39 and 40 might be allocated.

41. These different cases give rise to different interference restrictions on uses directly below channel 38. If the spectrum directly above channel 38 is unallocated or used for FDD-downlink, there is no restriction on uses directly below channel 38. However, if it is used for DVB-T or MMS, the separations with FDD-up and TDD exceed 8MHz and so are felt below channel 38. The various cases are summarised in the table below.

[7] This is a so-called set packing problem, which is NP-hard.

Table 3: Cases for channels 39 and 40

Use directly above channel 38	Implications below channel 38?	Number of cases for how channels 39 and 40 can be allocated
Unallocated	No	6
FDD downlink	No	4
DVB-T	Yes	2
MMS	Yes	2
TOTAL		**14**

42. We can now analyse the spectrum below channel 38 as a contiguous range using the same technique as for the upper sub-band, but in which there is a required separation on uses coming too close to channel 38. This varies according to the three cases in the table above (unallocated and FDD-downlink can be treated the same, as there is no restriction on spectrum below channel 38). Finally, we need to double the answer to allow for channel 38 itself either being allocated or unallocated. The overall results are given in Table 4 below.

Table 4: Upper bounds on the number of feasible usage vectors

	With lot categories as proposed by Ofcom	Without any subdivision of the usage rights into categories
Lower sub-band	181,614	5,886
Upper sub-band	16,335	669
TOTAL	2,966,664,690	3,937,734

43. With the lot structure proposed by Ofcom, the total number of feasible usage vectors is less than 3 billion. This means that problem should be soluble using a cluster of commodity computers and does not exceed the threshold where specialist techniques and equipment would be needed. However, this is still a very hard problem. At face value, this estimate gives a computational demand three orders of magnitude (i.e. over a thousand times) greater than the L-band auction.

44. In practice, this is a significant overestimate, as there are additional constraints on feasibility that we have not taken account of. In particular, once particular lot categories are fixed and potential conflicts with other lot categories identified, this will reduce the number of feasible usage vectors relative to our upper bound. However, refining these estimates is a significant task in itself, so we have yet to undertake this calculation.

45. A useful cross-check can be obtained by looking at a different, related problem. We can use the preceding analysis to count up the different cases of the number of lots in the usage right groups and ignore how these lots might be split amongst individual lot categories. In effect, this is the same as assuming that there was no differentiation within the five usage right groups other than:

- spectrum below channel 38, in channels 39 and 40, and in the upper sub-band being three separate categories within each usage right group; and
- channel 38 itself being split out as a separate lot.

46. Under these assumptions, the procedure followed above yields the exact number of feasible usage vectors, not an upper bound. This is because, within the three frequency ranges listed above, all lots are effectively moveable. Under these assumptions, there are about 4 million feasible usage vectors. Therefore, even if the lot structure were simplified greatly, this auction would still be a number of times more complicated than the L-band auction.

Contingencies if complexity increases

47. Given the toughness of the winner determination problem with the proposed lot structure, we should be looking to reduce rather than increase the number of lot categories from the 36 identified in this paper. Even with the current number, it might be found that moving to high performance computing solutions may not deliver results fast enough. Allowing for the fact that the estimate above is an upper bound, it looks as if even with some simplification of the lot structure, the problem is probably about two orders of magnitude harder than L-band.

48. Are there any alternatives? The first alternative is to reduce complexity by imposing more restrictions on how usage rights are allowed to fit together. For instance, requiring that rights appear in a particular order in the band reduces the number of possible outcomes. The difficulty is that this approach might eliminate some outcomes that are perfectly acceptable and efficient. However, practicality may simply require somewhat arbitrary rules to eliminate potential outcomes and reduce complexity. The idea would be to prune off unlike outcomes.

49. The second alternative is to use much faster algorithms that do not guarantee finding the overall optimal combination of winning bids. Such an approach is very common for tough real-world problems with similar characteristics to winner determination (so-called NP-hard problems). These methods try to achieve sequential improvements until such point that the rate of improvement achieved with each iteration falls below some threshold. This is the standard way of approaching problems such as transport route planning, scheduling, microchip layout design and so on.

50. The difficulty with using such methods for an auction is the unavoidable arbitrariness in the outcome selected. Although the total winning bids may be close to optimal, not being at a true global optimum could lead to the

winners and losers being different. Moreover, sequential adjustment algorithms may terminate a solution that depends both on where the algorithm is initialised from and exactly what adjustments are allowed. This would seem to create some potential for complaint from losers unless the methods to be used are very clearly laid out.

Annex 2: Algorithm for winner determination

Annex

This annex first outlines a simple algorithm for solving winner determination problems in package auctions with usage nominations (such as the recent L-band auction and the proposed DDR auction). We show that the number of feasible usage vectors is the key determinant of computational requirements for this algorithm. We then outline a method for bounding the number of feasible usage vectors.

1 A winner determination method

For simplicity, we assume that guard block requirements only arise as a result of the adjacency of dissimilar uses and ignore the possibility that different users with similar uses may require some separations.

1.1 Usage vectors and usage nominations

Let U be the set of usage categories. A usage vector $m : U \to \mathbb{N}$ describes how many lots there are in each category. Let M be the set of all feasible usage vectors, i.e. a list of the number of usage rights in each category that it is feasible to allocate. Conventionally, we include the possibility $0 \in U$, the category of unallocated lots. Let $\mathbf{0} \in M$ denote the trivial usage vector in which all available spectrum is unallocated, so that category 0 holds the maximum possible number of lots.

Suppose that there are I bidders and that any particular bidder $i \in \{1, \ldots, I\}$ makes J_i distinct package bids. A particular package bid $j \in \{1, \ldots, J_i\}$ made by bidder i is then described by:

- a bid amount b_{ij} for the overall package; and
- a *usage nomination* $m_{ij} \in M$ listing how many lots are required in each usage group.

Not all feasible usage vectors may necesarily be permissible usage nominations under the rules of auction. For example, it might be a requirement that any lots included in the package have the same usage nomination or that bids consist of a minimum number of lots. Conventionally, include also a 'null bid' of 0 for the trivial package $\mathbf{0}$ to represent the case of a bidder losing.

1.2 A dynamic programming algorithm

Dynamic programming is a practical method of solving small to medium scale combinatorial optimisation problems as it allows faster solutions than using general purpose integer programming algorithms. Define a value function $v_i(m)$ to be the greatest sum of winning bids that can be achieved if some lots have *already* been allocated according to the usage vector m and only the remaining unallocated lots are available for allocation to bidders $1, \ldots, i$. The overall solution to the WDP is given by computing $v_I(\mathbf{0})$.

The value functions satisfy a recurrent relationship - the Bellman equation - that expresses the optimal value for a given group of bidders in terms of the optimal value for smaller subsets of bidders. This relationship is directly analogous to that in the simpler case of a WDP without usage nominations. In particular, the optimal value for the first i bidders can be computed by considering which is the best package bid to accept from bidder i given the remaining lots available for allocating to the first $i-1$ bidders and subject to the requirement that the overall allocation is feasible (i.e. combining the package bid for bidder i with the existing usage vector must yield an feasible usage vector):

$$\begin{aligned} v_i(m) &= \max_j b_{ij} + v_{i-1}(m + m_{ij}) \\ &\text{s.t.} \quad m + m_{ij} \in M \end{aligned}$$

There is also an initial condition related to the first bidder

$$\begin{aligned} v_1(u) &= \max_j b_{1j} \\ &\text{s.t.} \quad m_{ij} \in M \end{aligned}$$

Notice that provided $\mathbf{0} \in M$ and we include a null bid (i.e. a bid of zero for no lots to represent the possibility of not being awarded anything) for each bidder, then there is always at least one feasible bid for any $m \in M$ that satisfies the conditions in the optimisation on the right-hand side of the Bellman equation. Therefore, under these conditions, the value functions are always defined and non-negative on M.

The optimal allocation can be obtained by taking the optimal bid choices given by the Bellman equations with an appropriate choice of the usage vector. Starting from bidder I and working downwards through the bidders, we find the optimal bid choice j_i^* for each bidder i given the usage vector m_i established by the allocation of lots to bidders $i+1, \ldots, I$, i.e.

$$\begin{aligned} j_i^* &= \arg\max_j b_{ij} + v_{i-1}(m_i + m_{ij}) \\ &\quad \text{s.t.} \quad m_i + m_{ij} \in M \\ m_{i-1} &= m_i + m_{ij} \end{aligned}$$

where $m_I = 0$. Where there are ties, there will be more than one way to define this sequence.

1.3 Computational requirements

The value functions are defined on the set M of all feasible usage vectors. Therefore, evaluation of the Bellman equation must be done once for each bidder and for each element of M. The number of operations within any particular evaluation of the Bellman equation depends on the number of bids for the bidder in question. This means that the number of operations required by this algorithm is linear in both the number of bids and in the size of M. When the number of bids is large, this is a very considerable advantage over direct search across all combinations of bids. Storage requirements for this algorithm are also proportional to the size of M. However, the size of M typically increases exponentially as we add additional categories.

2 Bounding the number of usage vectors

We now outline a general method for determining the number of feasible usage vectors.

2.1 Blocks and usage rights

Suppose that the sub-band in question consists of contiguous spectrum that breaks down into a set of B contiguous and equal sized *blocks*. These blocks form the basic unit of spectrum allocation. For DDR spectrum, these will be 1MHz blocks, as this is greatest common divisor of the different sizes of usage rights (i.e. 5MHz and 8MHz).

Bidders bid for *usage rights* that consist of multiple contiguous blocks. Bidders also nominate how the spectrum will be used, which determine compatibility with adjacent uses. Let U be the set of all possible usage right categories available in the sub-band. One of the possible uses will always be for a block to be unallocated, which we denote by $0 \in U$; this may be because the block is reserved as a guard block, or because when considering partial assignments of the spectrum this block has yet to be allocated to a bidder.

Given some usage right from category $u \in U$, let $n(u)$ be the number of contiguous blocks that a single lot contains. For unallocated blocks, $n(0) = 1$ by definition. At most $\lfloor |B| / n(u) \rfloor$ lots from category u can be assigned in this sub-band.[1] There may be many other limitations (to be considered in due course) that prevent this upper bound being achieved.

Let $b(u) \subseteq B$ be the set of blocks that can be used to form a usage right in category u. For example, if a usage right could be comprised of any of the blocks within the sub-band, we would simply have have $b(u) = B$. In other cases, then may be restrctions of which blocks could be used to form a usage right, in which case we would have $b(u) \subset B$. Typically $b(u)$ will be some interval contained within the sub-band corresponding to a frequency range.

[1] $\lfloor x \rfloor$ is the greatest integer smaller than x.

2.2 Usage vectors and patterns

A *usage vector* is simply a description of how *many* usage rights from each category are being allocated. Let $m : U \to \mathbb{N}$ be a particular usage vector; $m(u)$ is then the number of usage rights within category u allocated in the sub-band.

Notice that a usage vector is much simpler than a detailed sub-band plan. We do not specify where within the sub-band each usage right will be located. However, for a usage vector to be feasible, there must be at least one *pattern* of usage rights to particular locations within the sub-band that is consistent with all the various requirements that we impose, such as minimum separations amongst usage rights and any requirement that particular usage rights are located at particular places within the sub-band. It will often be the case that there are many patterns consistent with a given usage vector. In such a case, there may be rules restricting which patterns are possible as ultimate outcomes.

For the purposes of determining the winners at the end of the principal stage of a CCA what matters is the number of feasible usage vectors, not the number of possible assignments of usage rights (i.e. detailed band plans). The number of feasible usage vectors will be much smaller than the number of detailed band plans, so this is an important simplification in itself.

2.3 Guard block requirements

Where different uses come together, there may be some requirement to ensure a minimum separation between them. To model this, we make three simplifying assumptions:

- Usage rights can be grouped together into a number of groups (i.e. for the DDR spectrum there are five groups: DVB-T; MMS; TDD and FDD downlink and uplink);
- Usage rights within the same group all require the same number of blocks to construct one lot;
- Usage rights within the same group can be packed together contiguously without any guard block requirements between them;[2]
- The guard block requirements between two adjacent usage rights of different types is a function only of the groups within which the two adjacent usage rights fall (which implies more distantly located usage rights do not affect the guard block requirements); and
- Usage rights can be rearranged within the sub-band at will to minimise guard block requirements.

[2]This requirement is not very onerous and can be in fact be satisfied wherever the separation requirements between different usages is greater than between the usage rights of the same type. In such a case, we can simply redefine usage rights to include within the usage right any guard block required to separate similar usage rights; guard blocks are then only required when adjacent to a different usage right.

Let $G = \{g_1, g_2, \ldots, g_k\}$ be the non-trivial usage right groups which form a partition of the usage categories $U \setminus 0$. We are requiring that the number of blocks in each usage right is compatible with this partition, i.e. $n(u) = n(g)$ for all $u \in g$. Let $m(g) = \sum_{u \in g} m(u)$ be the total number of blocks allocated to usage group g under a usage vector m.

The first and second assumptions are innocent. The third presumes that all blocks within the sub-band are identical, so we can swap around usage rights any way we like to minimise guard block requirements. This is a useful simplification. However, in some cases there may be a need to keep some usage rights at particular locations or to respect some layout restrictions. In this case, not all possible rearrangements of usage rights within the sub-band may be possible. However, taking account of such restrictions on rearrangement will only increase the number of blocks that need to be retained as guard blocks.

Given the assumptions above, we know that to minimise overall guard block requirements within the sub-band, similar usage rights should be located together in contiguous blocks. This means that the total number of guard blocks that need to be left fallow in the sub-band is a function only of whether lots within a particular usage *group* are present or absent in the sub-band; once a usage group is present, the number of usage rights within the group or their precise breakdown into different types is irrelevant to the guard block requirement as this does not affect the nature of the boundaries between different usage rights.

This means that the total guard block requirement for the sub-band can be modelled in the following way. The non-trivial usage rights groups are $G \setminus 0$, where we exclude the 'unallocated' use. Let $\mathfrak{G} = 2^{G \setminus 0} \setminus \emptyset$ be the set of all possible non-empty subsets of the non-trivial usage groups. Then the guard block requirement can be written as a function $s : \mathfrak{G} \to \mathbb{N}$ where s is the number of blocks required for guard blocks if some subset of the usage groups are present in the sub-band.

A simple example makes this clear. Suppose that there are 10 blocks and two possible usage groups A and B that each use one block, but which need to be separated by at least two blocks. Then $\mathfrak{G} = \{\{A\}, \{B\}, \{A, B\}\}$, $s(\{A\}) = s(\{B\}) = 0$ and $s(\{A, B\}) = 2$.

Being able to express the guard block requirement as a function only of whether usage groups are present or absent is a massive simplification. Typically $\mathfrak{G}$ consists of a small number cases which we can analyse one-by-one. This number is *much* smaller than the number of possible usage patterns. In the case of the DDR spectrum, we can further eliminate any sets of usage groups that include FDD uplink without FDD downlink also being present.

Under the assumptions above, we can derive the number of total guard blocks required in the sub-band as a whole as a function of the usage groups present. Suppose that given two usage groups g_1 and g_2, there is a minimum separation requirement of $\sigma(g_1, g_2)$ blocks between this pair. Suppose further that if usage group g_1 is at the bottom end of sub-band (in terms of frequency) then a separation $\sigma_\ell(g_1)$ is needed from the adjacent out-of-band user below. Similarly if usage right g_2 is at the top end of sub-band (in terms of frequency) then a separation $\sigma_u(g_2)$ is needed from the adjacent out-of-band user above.[3] Then given some subset of non-trivial usage groups $\mathfrak{g} = \{g_{i_1}, \dots, g_{i_j}\} \subseteq \mathfrak{G}$ the minimum guard block requirement is then

$$s(\mathfrak{g}) = \min_{\pi \in \Pi_k} \left\{ \sigma_\ell(g_{i_{\pi(1)}}) + \sum_{r=1}^{j-1} \sigma\left(g_{i_{\pi(r)}}, g_{i_{\pi(r+1)}}\right) + \sigma_u(g_{i_{\pi(j)}}) \right\} \tag{1}$$

where Π_k is the set of all permutations of $\{1, \dots k\}$. Because all possible orderings of the different types of usage rights within the sub-band are possible, we can take which ever ordering is most efficient and minimises the overall need for guard blocks. This is the minimum number of guard blocks required to accommodate a particular set of usage groups simultaneously. This is only a lower bound on the number of guard blocks needed; in practice, some reorderings of the usage rights may not be possible so often the number of guard blocks required will be strictly greater than this.

2.4 Necessary conditions on feasible usage vectors

Given a particular set of usage groups being present, we know the maximum possible number of blocks that are available to allocate to non-trivial usages. We can cap the number of possible feasible usage vectors by considering all the different ways in which these blocks might be partitioned, first amongst the usage groups present, and then into individual usage categories.

A necessary condition for usage vector m to be feasible is that

$$|B| - s\left(\{g \in G \setminus 0 : m(g) > 0\}\right) \geq \sum_{u \in U} n(u)m(u) \tag{2}$$

This simply says that the number of available blocks is sufficient to meet the guard block requirement and then to allocate the required number of usage rights of each type (including any further unallocated blocks). The guard block requirement depends only on which usage groups are present in the usage pattern m, not the actual number of such rights.

We can now count the number of all such usage vectors satisfying condition (1). First, suppose that we know which usage rights are present or absent, that is we have some subset of usage rights $\mathfrak{g} \in \mathfrak{G}$. Then ask how many usage vectors there are satisfying condition (1) with exactly these usage rights being present in the sub-band. This is equal to the number of ways of dividing the available $|B| - s(\mathfrak{g})$ blocks amongst the usage rights present and then into individual usage categories.

[3] In the case of the lower DDR sub-band, we take B to be all those blocks below channel 38 and then run different scenarios for how the spectrum directly above channel 38 is used. Each scenario results in a different function σ_u depending how the spectrum above channel 38 is used. The number of cases in each scenario needs to be multiplied by the number of ways of allocating channel 38 itself (i.e. 2) and the number of ways of allocating channels 39 and 40 that give rise the presumed usage directly above channel 38. This is discussed in more detail in the main text.

To count these different ways of dividing up the blocks, we can first divide up the available blocks into usage groups, then subdivide these usage groups into usage categories. It is simplest to start with all the ways of dividing up r lots within a usage group exactly into $k = |g|$ categories within a usage group $g \subseteq U$. Define a partition function $P(r, k)$ to be the number of ways of doing this. This can be defined recursively by

$$P(r,1) = \begin{cases} 0 & \text{if } r = 0 \\ 1 & \text{otherwise} \end{cases} \qquad P(r,k) = \sum_{i=0}^{r} P(r-i, k-1).$$

We now count the number of ways of partitioning the blocks into usage groups and then into usage categories. Define a partition function

$$Q(K, \{(n_1, k_1), (n_2, k_2), \dots (n_j, k_j)\})$$

which is the number of ways of dividing up K objects into j groups such that:

- the number of objects in the ith group is a multiple of n_i;
- the number of objects in the ith group is at least n_i;
- there are no objects left over;
- the objects in the ith group are divided into k_i subgroups, with each subgroup containing a multiple of n_i objects.

We can evaluate Q recursively. If there is just one group, the number of objects must be divisible by the required size of 'lump' for that group otherwise we would have leftovers. In this case, we can form a certain number of lots within this single usage group that can then be subdivided over the available usage categories:

$$Q(K, \{n, k\}) = \begin{cases} P(K/n, k) & \text{if } K > 0 \text{ and } K \equiv 0 \bmod n \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

With more groups, Q can be expressed in terms of the corresponding expression with one less group by considering all the different possibilities for the number of objects that could be put into the first group and then subdivided over subgroups:

$$Q\left(K, \{(n_1, k_1), (n_2, k_2), \dots (n_j, k_j)\}\right) = \sum_{i=1}^{\lfloor K/n_1 \rfloor} Q\left(K - in_1, \{(n_2, k_2), \dots (n_j, k_j)\}\right) P(i, k_1) \tag{4}$$

These recursive formulae allow Q to be calculated on a computer.

Given this definition of the partition function, we obtain the following formula for the total number feasible usage patterns, $|M|$, on the assumption that all usage rights can be located anywhere within the sub-band

$$|M| = 1 + \sum_{\mathfrak{g} \in \mathfrak{G}} \sum_{k=1}^{|B| - s(\mathfrak{g})} Q(k, \{(n(g), |g|) : g \in \mathfrak{g}\}) \tag{5}$$

where $\{(n(g), |g|) : g \in \mathfrak{g}\}$ is a set of pairs each describing the size of a lot in a usage group g and the number of usage rights within that usage group. This can be interpreted in the following way. Take a non-empty subset of usage rights $\mathfrak{g}$. This requires $s(\mathfrak{g})$ guard blocks to be set aside as unallocated (at minimum). Of the available $|B| - s(\mathfrak{g})$ blocks suppose that k are allocated to usage rights and the remaining blocks are unallocated. There are $Q(k, \{(n(g), |g|) : g \in \mathfrak{g}\})$ ways of dividing k blocks into usage rights such that the usage groups present are exactly those assumed (i.e. $\mathfrak{g}$).

2.5 Further screening of infeasible usage vectors

The formula (5) gives a quick way of calculating (without enumeration) the exact number of feasible usage vectors for any given structure of usage rights and required separations provided that all usage rights can be placed anywhere within the band (i.e. $b(u) = B$ for all $u \in U$). However, this rests heavily on the assumption that usage rights can be rearranged within the sub-band without restriction. This assumption is not reasonable if certain usage rights need to go at particular locations (e.g. boundaries or within subsets of the available blocks). Also, the need to locate particular usage rights in particular points within the sub-band may interfere with our ability to move around usage rights to minimise the overall guard block requirement within the sub-band.

Nevertheless, the approach above is still useful as it provides a useful upper bound on the number of possible feasible usage vectors. If there are some additional restrictions on where usage rights can be located, this means that the expression (1) for the total guard block requirement may be too optimistic. However, it is still the case that the guard block requirement must be at least the number of blocks that would be required if all usage rights were re-arrangeable within the sub-band.

The second reason we might have an overestimate of the number of feasible usage pattens is that not all usage patterns may be compatible with at least one assignment of usage rights to locations within the sub-band if some usage rights need to go in specific locations. For instance, if there is only one place that a usage right can go, we can have at most one such usage right (e.g. a boundary lot). More generally, locating one usage right at a specific location within the sub-band may interfere with our ability to pack in the other usage rights.

The recursive methods above can easily be modified to enumerate all of the usage patterns described by formula (5). Some of these usage patterns may only be feasible under the assumption that all usage rights can be rearranged without restriction within the sub-band. We can then apply additional tests to knock out usage patterns that do not comply with our requirements.

One simple modification that we can make to the methods above is to take account of the limits that there might be on the number of usage rights of a particular type arising from requirements on their location within the sub-band. In particular, for any subset of blocks $\tilde{B} \subseteq B$, a feasible usage vector must satisfy the relationship

$$\sum_{b(u)\subseteq\tilde{B}} n(u)m(u) \leq \left|\tilde{B}\right|$$

i.e. that the total number of blocks allocated to any usage rights which must be located completely within the subset of blocks $\tilde{B}$ must not exceed the number of blocks contained within $\tilde{B}$. This is clearly true for the sub-band as a whole. Additional conditions are generated by considering strict subsets of the available blocks. Given that in practice the permissible ranges for locating lots in the DDR spectrum are frequency intervals within the band, we need only $\tilde{B}$ need only be an interval. Moreover, we can without loss of generality restrict to situations where the endpoints of $\tilde{B}$ are drawn from the endpoints of the permissible ranges from which the various categories of lots may be drawn.

This is again a necessary condition on feasible usage vectors and not a sufficiently one. However, it is simple to check by computer without needing to determine explicitly whether or not a certain number usage rights within each category can be packed within the available spectrum.