R. T. Chien
D. T. Tang

On Definitions of a Burst

One of the accepted definitions of a "burst" in coding theory is the following:*

"A burst of length b is a vector whose only nonzero components are among b successive components, the first and last of which are nonzero." (I)

This definition is quite adequate when one is interested in the theory of block codes of specified length whose capability of correcting or detecting burst errors is to be studied. However, difficulties can arise when it is used in the study of error statistics of a real channel in order to compare the merits of various burst-error-control codes.

Error data from real channels usually comes in the form of a sequence much longer than block lengths of practical codes. A typical error sequence in a binary system may have the following appearance:

$$\cdots 0 0 0 1 0 1 1 0 0 1 1 0 0 0 0 1 0 1 1 1 1 0 0 \cdots \quad (1)$$

There seems to be no simple way of using definition (I) to describe such an error sequence. The difficulty lies in two facts: first, that the bursts have nonunique representations of different lengths and, second, that the block length of the code to be used is yet undetermined.

To circumvent this difficulty, Definition (II) may be used:†

"A burst of length b is a sequence of b digits, the first digit of which is nonzero." (II)

Since the last digit of the burst is no longer required to be nonzero, the total number of bursts with a specific length can be readily determined in a given error sequence. There is no confusion due to mixed burst lengths. For instance, the sequence in (1) contains 10 bursts of length one, or 6 bursts of length two, or 5 bursts of length three, or four bursts of lengths four and five. A table of total numbers of bursts of different lengths is easily made available for the selection of codes.

The purpose of this note is to show that Definition (II) is quite adequate for general use in the treatment of error-correcting codes, and that it offers advantages when applied to the error analysis. Bounds on the performance of burst-error-correcting codes can be obtained from the simple counts of such defined bursts of specific lengths in a given error sequence.

With the new definition of (II), a linear burst-error-correcting code may be defined as follows:

A linear code is said to be capable of correcting a single burst of length b if all distinct error patterns each containing a burst of length b belong to distinct cosets.

In determining whether a particular error pattern is correctable by a burst-b correcting code, all that one must do is to see whether the error pattern contains a single burst of length b. When an error sequence such as that in (1) is given, a sufficient condition for a burst-b correcting block code of length n to correct all the errors in the error sequence is to have

$$N(b) = N(n + 2b - 2), \quad (2)$$

where $N(i)$ is the total number of bursts of length i in the sequence. The sufficiency of (2) can be verified as follows: Since Eq. (2) implies that no burst of length $(n + 2b - 2)$ in the error sequence contains all the 1's in any two adjacent bursts of length b, a minimum of $(n + 2b - 1)$ digits is necessary to cover all the 1's in any adjacent b-bursts. It follows that any two b-bursts must be separated by at least $(n - 1)$ error-free digits. Such an error sequence is clearly correctable by a b-burst-correcting code of length n. If the equality of Eq. (2) does not hold, then the difference $N(n + 2b - 2) - N(b)$ can be used as an upper bound to the total number of uncorrectable blocks.

* See p. 60 of "Error Correcting Codes" by W. W. Peterson. The MIT Press and John Wiley and Sons, New York, 1961.
† Definition II has been used in error analysis experiments on telephone lines. See A. A. Alexander, R. M. Gryb, and D. W. Nast, "Capabilities of the Telephone Network for Data Transmission," *Bell System Tech. J.* **39**, 3 (May 1960).

Since Eq. (2) guarantees that any two b-bursts are separated by at least $(n-1)$ error-free digits, one sees that Eq. (2) is also a sufficient condition for a recurrent code to correct a b-burst in any span of n digits, or to correct b-bursts with a minimum error-free guard space of $(n-1)$. Furthermore, a necessary condition can easily be shown to be:

$$N(n) = N(b). \qquad (3)$$

Such conditions are useful in estimating the performance of burst-error-correcting codes and in reducing unnecessary computations if detailed code simulation is to be carried out next.

Three examples follow:

For a burst-2 correcting block code of length $n = 7$, the error sequence

$$\cdots 0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,1\,1\,0\,0\,0\,0\,0\,0$$
$$0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0 \cdots \qquad (4)$$

is correctable since $N(2) = N(9) = 3$.

The error sequence

$$\cdots 0 \,|\, 0\,0\,0\,1\,0\,0\,0 \,|\, 0\,0\,0\,0\,0\,1\,1$$
$$|\, 1\,1\,0\,0\,0\,0\,0 \,|\, 1\,0\,0\,0\,0\,0\,0 \,|\, \cdots \qquad (5)$$

is correctable only with respect to the block partition shown. Here $N(2) = 4$, and $N(9) = 3$.

The error sequence

$$\cdots 0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1\,1$$
$$0\,0\,0\,0\,1\,0\,0\,0\,0\,0\,0\,0 \cdots \qquad (6)$$

with $N(2) = 4$ and $N(9) = 2$ cannot be completely corrected regardless of the block partition.

For a recurrent code which corrects a burst-2 error in any span of seven digits, or which corrects 2-bursts with a minimum error-free guard space of six digits, only the sequence in (4) is correctable. Sequences in (5) and (6) both violate condition (3).

Received April 16, 1965