

INTERNATIONAL TELECOMMUNICATION UNION

**TELECOMMUNICATION  
STANDARDIZATION SECTOR**

STUDY PERIOD 2009-2012

**STUDY GROUP 16  
TD 355 (PLEN/16)**

**English only**

**Original: English**

---

**Question(s):** 16/16

Geneva, 14 - 25 March 2011

**TEMPORARY DOCUMENT**

**Source:** Editor G.160

**Title:** Approval: ITU-T G.160 Amd.2 “Voice enhancement devices: Revised Appendix II – Objective measures for the characterization of the basic functioning of noise reduction algorithms” (New)

---

This TD contains the text of G.160 Amendment 2. Since the changes only affect *non-normative* parts of the Recommendation, the text is put forward for Approval at this SG 16 meeting.

The highlighted text is the revision with respect to Amendment 1 of G.160 (11/09). The electronic attachment (G160Amendment2ElectronicAttachment.zip) is also included.

---

**Contact:** Dominic Ho  
Ericsson  
Canada

Tel: +1 573 823 6188  
Fax:  
Email: [hodmissouri@yahoo.com](mailto:hodmissouri@yahoo.com)

<p><b>Attention:</b> This is not a publication made available to the public, but <b>an internal ITU-T Document</b> intended only for use by the Member States of ITU, by ITU-T Sector Members and Associates, and their respective staff and collaborators in their ITU related work. It shall not be made available to, and used by, any other persons or entities without the prior written consent of ITU-T.</p>
---

## **Draft new Recommendation ITU-T G.160 Amendment 2**

### **Voice enhancement devices: Revised Appendix II – Objective measures for the characterization of the basic functioning of noise reduction algorithms**

#### **Summary**

Appendix II of Recommendation ITU-T G.160 presents an objective methodology for characterizing the basic effect of noise reduction (NR) methods. Three objective measures are specified for characterizing NR solutions. Recommendation target value ranges in specified conditions are given for the measures to serve as basic guidelines for proper functioning of NR methods. Amendment 1 of Recommendation ITU-T G.160 brings corrections and clarifications to the text of Appendix II. [Amendment 2 sees the addition of ANSI-C source code and test vectors as an electronic attachment. Some changes have also been made to improve the accuracy of the SNRI metric when the output SNR exceeds approximately 25 dB.](#)

#### **Electronic attachment:**

- ITU-T G.160 Amendment 2 ANSI-C source code and test vectors

[http://itu.int/md/dologin\\_md.asp?id=T09-SG16-110314-TD-PLN-0355!A1!ZIP-E&type=mitems](http://itu.int/md/dologin_md.asp?id=T09-SG16-110314-TD-PLN-0355!A1!ZIP-E&type=mitems)

**(Full text available only in the electronic version)**

## CONTENTS

### Page

Appendix II	Objective measures for the characterization of the basic functioning of noise reduction algorithms .....	3
II.1	Introduction .....	3
II.2	Notations.....	3
II.3	Test signals .....	4
II.4	Objective measures for characterization of NR algorithm effect .....	5
II.4.1	Frame categorization for speech and speech pause classes.....	5
II.4.2	Assessment of SNR improvement – <i>SNRI</i> .....	8
II.4.3	Assessment of total noise level reduction – <i>TNLR</i> .....	9
II.4.4	Comparison of <i>SNRI</i> and <i>NPLR</i> – <i>DSN</i> .....	10
II.4.5	Block diagram presentation of proposed measures .....	10
II.5	Recommendation values for <i>SNRI</i> , <i>TNLR</i> and <i>DSN</i> .....	12
II.6	Reference software .....	12

## Appendix II

### Objective measures for the characterization of the basic functioning of noise reduction algorithms

(This appendix does not form an integral part of this Recommendation)

#### II.1 Introduction

This appendix presents an objective methodology for characterizing the basic effect of noise reduction (NR) methods. Three objective measures are specified for characterizing NR solutions. Recommendation target value ranges in specified conditions are given for the measures to serve as basic guidelines for proper functioning of NR methods.

The methodology monitors the basic functioning of an NR solution in terms of signal-to-noise ratio improvement (*SNRI*) and total noise level reduction (*TNLR*). While *SNRI* is measured during speech activity, focusing on the effect of NR on speech signals, *TNLR* estimates the overall level of noise reduction, experienced both during speech and speech pauses. In addition, a delta measurement (*DSN*) is computed to reveal speech attenuation or undesired speech amplification caused by an NR solution.

The proposed methodology is a further development of one included in [b-ETSI TS 126 077] and [b-TIA/EIA-IS-853]. A detailed description of the earlier methodology is incorporated in the named specifications.

NOTE 1 – "Noise suppression" in the named standard specifications refers to methods that reduce the background noise impact in the desired speech signal, with no limitation as to their capability of reducing the noise level either only outside of or also during speech activity. The terms "noise suppression" and "noise reduction" are used interchangeably in this appendix.

NOTE 2 – It should be noted that these objective measures of noise reduction performance do not measure speech quality or voice distortion, which should be taken into account when assessing the overall performance of a noise suppressor.

NOTE 3 – The accuracy of the metrics calculated using the method described in this Appendix can vary depending on the type of noise reduction algorithm being tested. For certain noise suppression systems, the SNRI metric shows a bias to the real SNRI. A theoretical study, confirmed by simulations, indicates negligible bias in SNRI metric at output SNR below 25dB. The bias remains below 3dB until 35dB output SNR but it may increase further with higher output SNR.

#### II.2 Notations

The following notations are used in this appendix:

- The operator  $NR(\cdot)$  corresponds to applying the NR algorithm on the input speech.
- The clean speech original signals are referred to as  $s_i$ ,  $i = 1$  to  $I$ , where  $I$  is the total number of speech files used.
- The noise original signals are referred to as  $n_j$ ,  $j = 1$  to  $J$ , where  $J$  is the total number of noise files used.
- The noisy speech test signals are referred to as  $d_{ij} = s_i + \beta_{ij}(\text{SNR}) n_j$ ,  $i = 1$  to  $I$ ,  $j = 1$  to  $J$ , where  $d_{ij}$  is built by adding  $s_i$  and  $n_j$  with a pre-specified SNR as presented below.
- The processed signal is referred to as  $y_{ij} = NR(d_{ij})$ .
- The reference signal in the calculations shall be the noisy speech test signal  $d_{ij}$  itself.
- The notation  $\text{Log}(\cdot)$  indicates the decimal logarithm.

- $\beta_{ij}(\text{SNR})$  is the scaling factor to be applied to the background noise signal  $n_i$  in order to have a ratio **SNR** (in dB) between the clean speech signal  $s_i$  and  $n_j$ . The use of the scaling factor is explained below in clause II.3, in conjunction with the generation of test samples.
- Frames are classified by their average speech power (high, medium, low); these classes are named  $\{k_{\text{sph}}\}$ ,  $\{k_{\text{spm}}\}$ ,  $\{k_{\text{spl}}\}$ .
- Frames containing only noise form a separate class, named  $\{k_{\text{nse}}\}$ .
- The determination of which frames contain active speech is to be carried out with reference to [b-ITU-T P.56] and is related to the classification of the frames into the presented categories, or *speech power classes* which is explained below.

### II.3 Test signals

The test material should manifest at least the following extent:

- Clean speech utterance sequences: 6 utterances from 4 speakers – 2 male and 2 female – totalling 24 utterances
- Noise sequences:
  - Car interior noise, 120 km/h, fairly constant power level
  - Street noise, slowly varying power level
  - Babble noise

Special care should be taken to ensure that the original samples fulfil the following requirements:

- The clean speech signals are of a relatively constant average (within sample, where 'sample' refers to a file containing one or more utterances) power level and have sufficient short speech pause frames.
- ~~The noise signals are of a short-time stationary nature with no rapid changes in the power level and no speech-like components~~

The test signals should cover the following background noise and SNR conditions:

- Car noise at 6 dB and, 12 dB ~~and 18 dB~~ SNR
- Street noise at 6 dB, ~~12 dB~~ and 12 dB SNR
- Babble noise at 6 dB and 12 dB SNR

Where the amount of NR is not expected to exceed 17dB, an input SNR of 18 dB should be used in addition to 6 dB and 12 dB.

The samples should be digitally filtered before NR processing by an appropriate filter to become representative of a real ~~networkcellular~~ system frequency response. This filtering should be carried out before the scaling of the samples to be discussed below. After pre-filtering and normalization of the active speech level to -26dBov, the clean speech signals should have a noise floor lower than -61dBov for every frame in the interval between utterances. This is to ensure that NPLR and SNRI calculations are based only on the short pause frames that are amongst speech activity (within a speech sentence) excluding all frames during the long speech pauses (between utterances).

NOTE – Appropriate digital filters to make the test material representative of a real cellular mobile station frequency response, which are available as part of the G.191 software tool library, are the following:

- The modified IRS filter according to [b-ITU-T P.48].
- The MSIN (mobile station in) filter. The transfer function of the MSIN filter has been defined based on measurements by British Telecom. Even though the MSIN filter attempts mainly to describe the spectral modifications of background noise, in the ETSI (European Telecommunications Standards

Institute) STC SMG11#8 Plenary (Jan 1999; Helsinki, Finland), it has been discussed that this filter may also be applied to the speech signal.

The test samples corresponding to the named SNR conditions shall be generated according to the following procedure:

- The clean speech material shall be scaled to the active speech level  $-26$  dBov with the [b-ITU-T P.56] speech voltmeter, one file at a time, each file including a sequence of one (to four) utterance(s) from one speaker.
- A silence period of 2 s is inserted at the beginning of each of the resulting files to make up augmented clean speech files.
- Within each noise type and SNR condition, a noise sequence is selected for every augmented clean speech utterance file, each with the same length as the corresponding speech file, and each noise sequence is stored in a separate file.
- Each of the noise sequences is scaled to a dBov level leading to the SNR condition corresponding to the  $\beta_{ij}(\text{SNR})$  value in each of the test cases by applying the RMS level based scaling according to [b-ITU-T P.56].
- Augmented clean speech and noise files are then added sample by sample (overload values being clipped) to create the needed set of noisy speech files with predefined SNR, sample-rate, bits/sample, file-length.
- Input and output signals in all metric calculations must be time synchronized with the clean speech signal.

## II.4 Objective measures for characterization of NR algorithm effect

### II.4.1 Frame categorization for speech and speech pause classes

The objective metric for measuring SNR improvement due to an NR algorithm, or *SNRI*, that is presented below, is calculated separately in three groups of speech signal frames that represent power-gated constituents of active speech signal. These groups of frames are called *speech power classes* or just *power classes* in the following. Hence, the *SNRI* measure is calculated separately in frames of high, medium and low power. These categories are used to characterize the effect of the NR processing on speech, allowing the distinction of the effect on strong, medium and weak speech. In addition to calculating the SNR improvement separately on the three categories, they are used to form an aggregate measure.

In addition to the division into three frame classes containing speech activity, a fourth frame class is needed to be able to calculate the power levels in the noise periods lying amongst speech activity. These power levels are used both in the calculation of the *SNRI* metrics and in the calculation of the overall change in noise power level (*TNLR*) due to NR processing.

A frame length of 10 ms, i.e., 80 samples (for speech sampled at 8 kHz), is used in the analysis since it has been found the most efficient to describe changes in the signal caused by NR processing.

To determine which frames belong to high, medium and low power classes of active speech and which present pauses in the speech activity (noise only), the active speech level (in dB) ( $sp\_lvl$ ) of the noise free speech  $s_i(n)$  is first determined according to [b-ITU-T P.56]. Thereafter, the frames are classified into the four classes based upon a comparison of the power in each frame to predefined threshold values relative to the active speech level. As a result, four number sequences are obtained to represent the three speech classes and the speech pause class, respectively:  $\{k_{sph}\}$ ,  $\{k_{spm}\}$ ,  $\{k_{spl}\}$ ,  $\{k_{nse}\}$ .

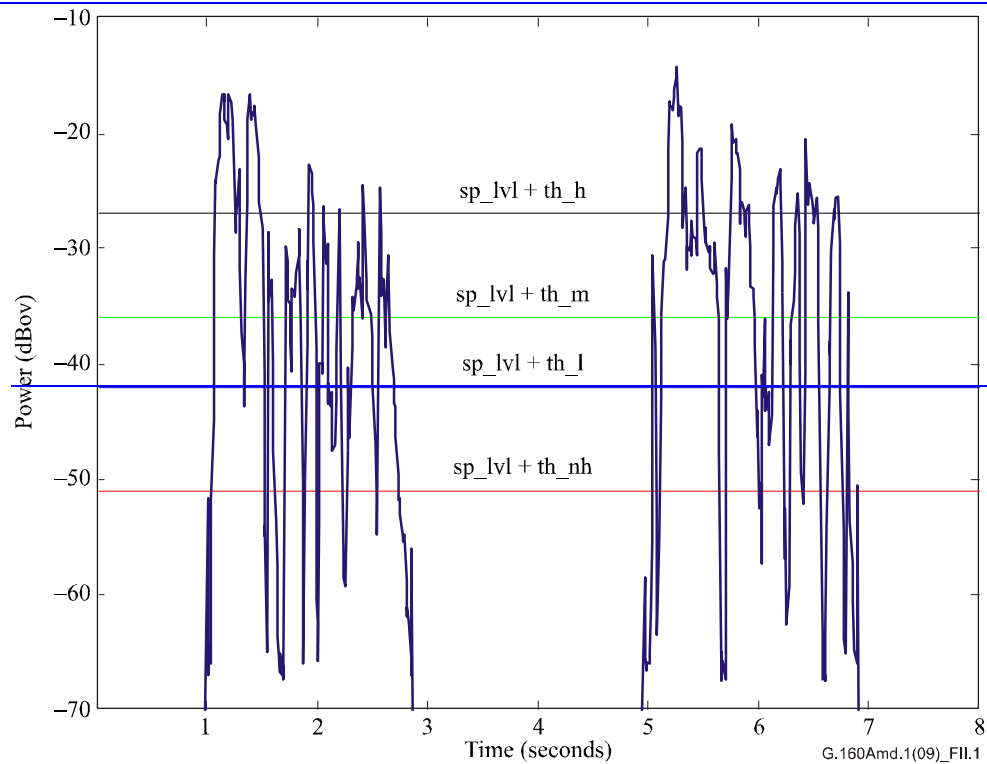
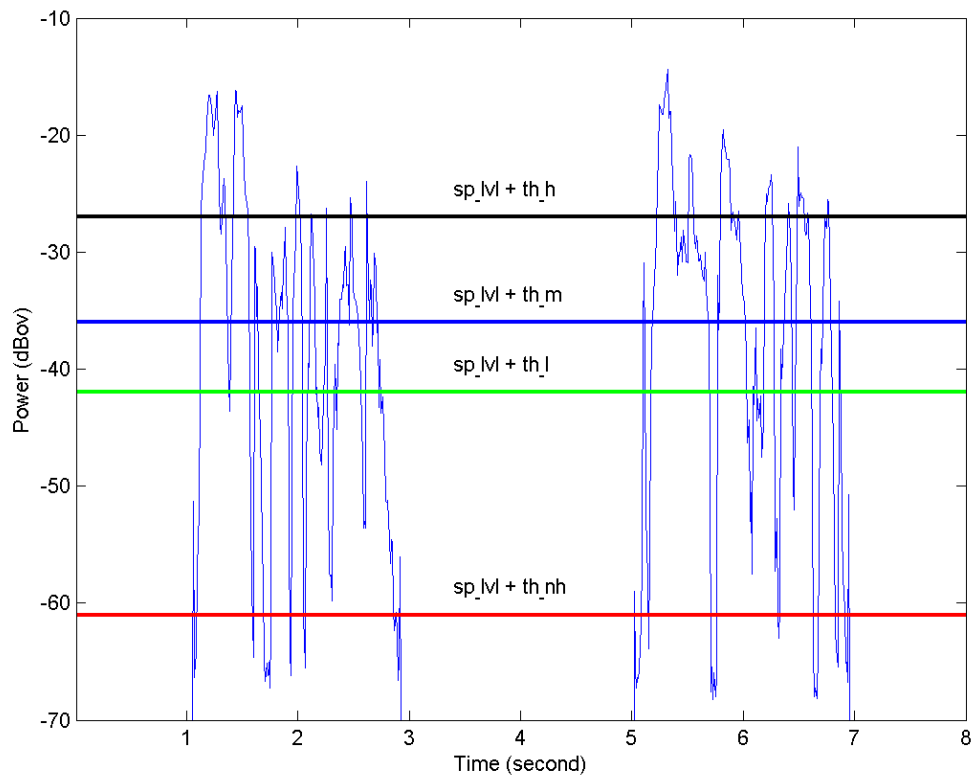
The threshold values to be related to the active speech level are presented in Table II.1. To be able to apply the threshold values given here, the speech should be first normalized at  $-26$  dBov using the [b-ITU-T P.56] speech voltmeter. Figure II.1 presents an example of the classification of a clean speech sample into power classes.

**Table II.1 – Power threshold values for speech power classes**

Threshold	Explanation	Value
th_h	Lower bound for high speech power class	$-1$ dB
th_m	Lower bound for medium speech power class	$-10$ dB
th_l	Lower bound for low speech power class	$-16$ dB
th_nh	Higher bound for speech pause class	$-32.5$ dB
Dlp_min	Minimum duration for long speech pause class	400 ms

The following notes are made concerning the formulation of the speech power classes:

- All frames with power level lower than the higher bound for speech pause class,  $sp\_lvl + th\_nh$ , are used in the computation of  $TNLR$  as this should indicate the overall reduction in noise level experienced both during short speech pauses amongst speech activity and long speech pauses when no speech is immediately continuing. On the other hand, only those frames in speech pauses of duration less than the minimum duration for long speech pauses,  $Dlp\_min$ , are applied to the computation of both  $SNRI$  and  $DSN$ . This exclusion of long pauses from the analysis makes  $SNRI$  and  $DSN$  concentrate on those noise frames among or close to speech activity.
- In low SNR conditions, the noise power level may be higher than the lower bound of some of the speech power classes. However, even in this case, the information of the effect on the low power portions of speech may be informative.



**Figure II.1 – Power in frames of clean speech as a function of time**  
**Classification into power classes**



## II.4.2 Assessment of SNR improvement – *SNRI*

NOTE – See NOTE 3 in Clause II.1 for guidance on the accuracy of SNRI.

The SNR improvement measure, *SNRI*, measures the SNR improvement achieved by the NR algorithm. As said above, SNR improvement is calculated in three *speech power classes* to obtain an evaluation of the effect separately for strong, medium and weak speech. An aggregate measure is obtained in addition via weighted averaging.

The calculation is here presented for the high power speech class. A similar procedure is used to obtain the respective figures for the medium and low power speech classes.

For each background noise condition  $j$  and for each speech sample  $i$ , construct a noisy speech input signal  $d_{ij}$  as follows:

$$d_{ij}(n) = \beta_{ij} n_j(n) + s_i(n),$$

where  $\beta_{ij}$  depends on the SNR condition according to the procedure described above. Now, the noise suppressed output speech signal can be denoted as:

$$y_{ij} = \text{NR}(d_{ij})$$

In this way, the SNR of the output and input speech signals can be expressed as:

$$SNR_{out\_h_{ij}} = 10 \cdot \text{Log} \left\{ \max \left[ \varepsilon, \frac{10^{\frac{1}{K_{sph}} \sum_{l=1}^{K_{sph}} \text{Log} \left\{ \max_n \left[ \xi, \sum_n y_{ij}^2(l, n) \right] \right\}}}{10^{\frac{1}{K_{nse}} \sum_{m=1}^{K_{nse}} \text{Log} \left\{ \max_p \left[ \xi, \sum_p y_{ij}^2(m, p) \right] \right\}}} - 1 \right] \right\} \quad (\text{II.1})$$

$$SNR_{in\_h_{ij}} = 10 \cdot \text{Log} \left\{ \max \left[ \varepsilon, \frac{10^{\frac{1}{K_{sph}} \sum_{l=1}^{K_{sph}} \text{Log} \left\{ \max_n \left[ \xi, \sum_n d_{ij}^2(l, n) \right] \right\}}}{10^{\frac{1}{K_{nse}} \sum_{m=1}^{K_{nse}} \text{Log} \left\{ \max_p \left[ \xi, \sum_p d_{ij}^2(m, p) \right] \right\}}} - 1 \right] \right\} \quad (\text{II.2})$$

$$SNRI_{h_{ij}} = SNR_{out\_h_{ij}} - SNR_{in\_h_{ij}}, \quad (\text{II.3})$$

where

$K_{sph}$  is the total number of frames containing speech of high power;

$K_{nse}$  is the total number of frames containing noise during short speech pauses;

$\xi > 0$  is a constant that should be set at  $8 \cdot 10^{-8}$ ;

$\varepsilon > 0$  is a constant that should be set at 0.0631 (equivalent to -12 dB).

The summation with respect to indices  $n$  and  $p$  is carried out in frames of 80 samples.  $SNRI_{m_{ij}}$  and  $SNRI_{l_{ij}}$  are computed correspondingly for medium and low power frames.

Finally, the aggregate measure for signal-to-noise ratio improvement is computed as below.

$$SNRI_{ij} = \frac{1}{K_{sph} + K_{spm} + K_{spl}} (K_{sph} \cdot SNRI_{h_{ij}} + K_{spm} SNRI_{m_{ij}} + K_{spl} SNRI_{l_{ij}}) \quad (\text{II.4})$$

$$SNRI_j = \frac{1}{I} \sum_{i=1}^I SNRI_{ij} \quad (\text{II.5})$$

$$SNRI = \frac{1}{J} \sum_{j=1}^J SNRI_j \quad (\text{II.6})$$

In addition, measures for the SNR improvement in the high, medium and low power speech classes (SNRI\_h, SNRI\_m, SNRI\_l, respectively) shall be recorded based on the following formulae:

$$SNRI\_h = \frac{1}{J} \sum_{j=1}^J SNRI\_h_j = \frac{1}{J} \sum_{j=1}^J \frac{1}{I} \sum_{i=1}^I SNRI\_h_{ij} \quad (II.7)$$

$$SNRI\_m = \frac{1}{J} \sum_{j=1}^J SNRI\_m_j = \frac{1}{J} \sum_{j=1}^J \frac{1}{I} \sum_{i=1}^I SNRI\_m_{ij} \quad (II.8)$$

$$SNRI\_l = \frac{1}{J} \sum_{j=1}^J SNRI\_l_j = \frac{1}{J} \sum_{j=1}^J \frac{1}{I} \sum_{i=1}^I SNRI\_l_{ij} \quad (II.9)$$

It is, in addition, informative to record separately the noise type specific SNR improvement measures, namely,  $SNRI\_h_j$ ,  $SNRI\_l_j$ ,  $SNRI\_m_j$  and  $SNRI_j$  for each background noise condition  $j$ .

#### II.4.3 Assessment of total noise level reduction – *TNLR*

The total noise level reduction measure, or *TNLR*, relates to the capability of the NR method to attenuate the background noise level measured during both speech activity and speech pauses. As the number of frames in the speech pause class during speech activity is typically relatively small compared to the number of frames during long speech pauses, *TNLR* mainly measures the capability of an NR to reduce noise during long speech pauses.

The *TNLR* measure is calculated as follows:

For each background noise condition  $j$  and for each speech sample  $i$ , construct a noisy input signal  $d_{ij}$  as follows:

$$d_{ij}(n) = \beta_{ij} n_j(n) + s_i(n),$$

where  $\beta_{ij}$  depends on the SNR condition according to the procedure described above. The noise suppressed output speech signal can be denoted as:

$$y_{ij} = \text{NR}(d_{ij})$$

In this way, the total noise level reduction measure can be expressed as:

$$TNLR_{ij} = \frac{1}{K_{pse}} 10 \cdot \sum_{m=1}^{K_{pse}} \left[ \text{Log} \left\{ \max \left( \xi, \sum_q d_{ij}^2(m, q) \right) \right\} - \text{Log} \left\{ \max \left( \xi, \sum_q y_{ij}^2(m, q) \right) \right\} \right], \quad (II.10)$$

where

$K_{pse}$  is total number of noise frames during speech pauses (frame power of clean speech signal  $< \text{sp\_lvl} + \text{th\_nh}$ ) with active noise (frame power of noisy input signal  $> \text{cft\_lvl}$ );

$\xi > 0$  is a constant that should be set at  $8 \cdot 10^{-8}$

$\text{cft\_lvl}$  is a comfort noise level constant that will default to  $-48$  dB

$$TNLR_j = \frac{1}{I} \sum_{i=1}^I TNLR_{ij} \quad (II.11)$$

$$TNLR = \frac{1}{J} \sum_{j=1}^J TNLR_j \quad (II.12)$$

The summation with respect to index  $q$  is carried out in noise frames of 80 samples. The index,  $q$ , relates to noise frames with frame power less than the higher bound for speech pause class,  $th\_nh$  (see Table II.1).

Furthermore, it is informative to record separately the noise type specific  $TNLR$  measures, or  $TNLR_j$ , for each background noise condition  $j$ .

#### II.4.4 Comparison of $SNRI$ and $NPLR - DSN$

One of the targets in noise suppression is to maintain the power level of the speech signal so as not to attenuate the level of the speech signal together with the noise signal in the NR processing. Furthermore, an improvement in the SNR of a noisy speech signal may be obtained in a simple amplification of the speech signal, the amplification being typically only applied to the high-energy portions of the signal. As both the attenuation and amplification of speech are not desirable in NR, they should be detected.

Both the attenuation and amplification of a speech signal due to an NR can be measured by studying the balance between the SNR improvement and the noise level reduction obtained during speech activity. For this purpose, noise power level reduction, or  $NPLR$ , is measured during short speech pauses lying amongst speech activity to study the capability of an NR to reduce noise from speech. This measure is computed otherwise as  $TNLR$ , i.e., according to equations II.10, II.11 and II.12, except that the computation is only carried out with frames classified as short pause frame. Here a frame of speech pause class is classified as a short pause frame if its duration is less than the minimum duration of long pause,  $Dlp\_min$ , which is set at 400 ms. In this way,  $NPLR$  provides a counterpart for  $SNRI$ , these metrics forming together the basis for the evaluation of the balance. As  $NPLR$  is used as an internal variable, it is not reported.

$SNRI$ -to- $NPLR$  difference ( $DSN$ ), comprising a comparison of the  $SNRI$  and  $NPLR$  measures, is therefore proposed as a measure to acquire an indication of possible speech attenuation or speech amplification produced by the tested NR method and is formulated as:

$$DSN = SNRI - NPLR \quad (II.13)$$

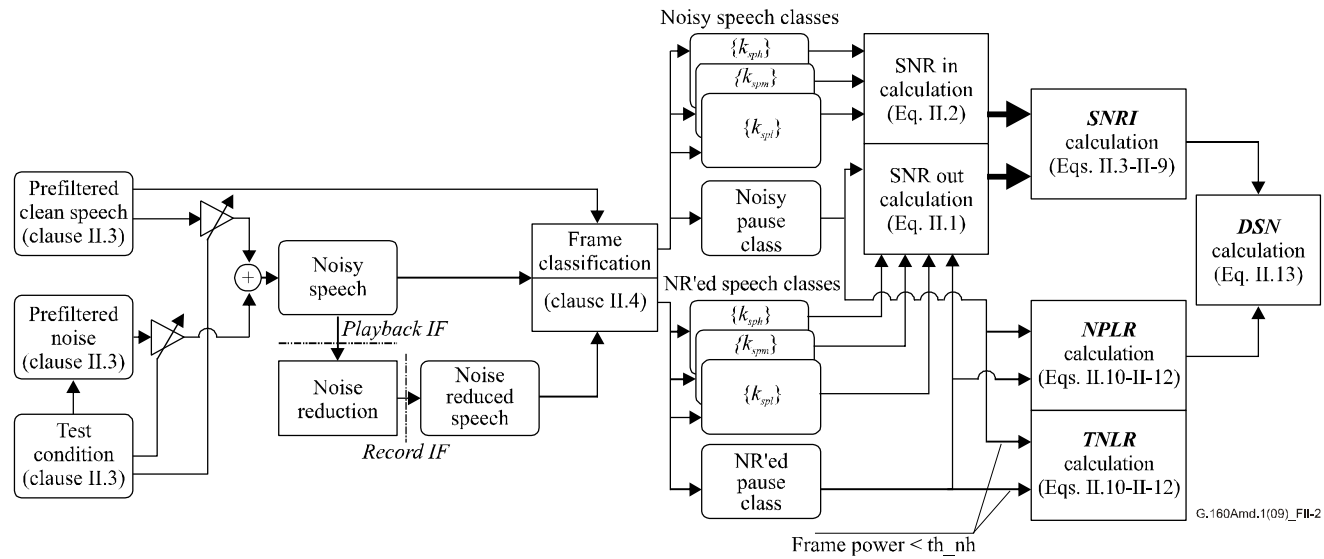
Typically the  $DSN$  quantity should get values close to zero. If the  $NPLR$  parameter assumes clearly higher values than  $SNRI$ , making  $DSN$  clearly negative, the NR solution turns out to produce speech level attenuation. On the other hand, if  $DSN$  becomes clearly positive, it indicates the speech signal has been amplified, which contributes to the SNR improvement partially or wholly if no reduction in noise level.

It is informative to record separately the noise type specific  $DSN$  measures, or  $DSN_j$ , for each background noise condition  $j$ .

#### II.4.5 Block diagram presentation of proposed measures

A block diagram of the calculation for the proposed metrics is provided in Figure II.2. It should be noted that an appropriate prefiltering method (see clause II.3) should be used in each studied case to make the speech data representative of real network conditions.

If the measurement of the NR device ("noise reduction") is made in a network system implementation, the interfaces for playing back the NR input signal ("*Playback IF*") and recording ("*Record IF*") the NR output signal can be utilized. The rest of the processing and calculation stages can be conducted off-line.



**Figure II.2 – Block diagram of calculation of *SNRI*, *TNLR*, *NPLR* and *DSN* measures for noise reduction algorithms**

## II.5 Recommendation values for *SNRI*, *TNLR* and *DSN*

The objective measures of signal-to-noise ratio improvement (*SNRI*), total noise level reduction (*TNLR*), and *DSN* defined above are to be used to characterize the performance of the NR solution. Objectives are defined for these measures in Table II.2. These measures can be used to provide an indication of the benefit produced by an NR solution but should be interpreted as informative.

It is noted that NR solutions may function in different ways in varying usage conditions in terms of the objective metrics presented above. Therefore, for *SNRI*, *TNLR* and *DSN*, the performance objectives are defined for average values of these measures over all test conditions.

**Table II.2 – NR algorithm performance targets in *SNRI*, *TNLR* and *DSN* metrics**

Objective quality measure/test condition	Performance objective
<i>SNRI</i> <del>Assessment: To be evaluated according to clause II.4 using a predefined set of material comprising speech mixed with stationary car noise and street noise in the SNR conditions defined in clause II.3.</del>	$SNRI \geq 4$ dB as an average over all test conditions
<i>TNLR</i> <del>Assessment: To be evaluated according to clause II.4 using a predefined set of material comprising speech mixed with stationary car and street noise in the SNR conditions defined in clause II.3.</del>	$TNLR \geq 5$ dB as an average over all test conditions
<i>DSN</i> <del>Assessment: To be evaluated according to clause II.4 using a predefined set of material comprising speech mixed with stationary car noise and street noise in the SNR conditions defined in clause II.3.</del>	$-4 \text{ dB} \leq DSN \leq 3 \text{ dB}$ as an average over all test conditions

NOTE – In Table II.2 above, all of the metrics should be evaluated according to clause II.4 using a predefined set of material comprising speech mixed with stationary car noise, street noise and babble noise in the SNR conditions defined in clause II.3.

Meeting the presented performance objectives can be regarded as indicating that the NR method under study is capable of producing significant benefit over not using noise reduction. On the other hand, different tuning of an NR algorithm design may be beneficial in different usage scenarios, e.g., with regard to the number of speech transcoding and/or voice enhancement processing stages in the speech path. Hence, no strict requirements on the NR performance in the specified metrics are set.

## II.6 Reference software

This Appendix includes an electronic attachment containing ANSI-C source code and a set of test vectors. Table II.3 lists the files that comprise the electronic attachment.

NOTE – The test vectors provided are examples of input signals that can be used with the metrics described in this Appendix. Other input signals can be used, providing the guidance in section II.3 is followed.

**Table II.3 – Organization of the electronic attachment**

<b><u>File/directory name</u></b>	<b><u>Description</u></b>
<a href="#">Readme.txt</a>	<a href="#">A brief introduction to the package</a>
<a href="#">g160app2.c</a>	<a href="#">Main program</a>
<a href="#">sv-p56.c</a>	<a href="#">Function for ITU-T P.56 active speech level meter</a>
<a href="#">sv-p56.h</a>	<a href="#">Header file for sv-p56.c</a>
<a href="#">g160app2.exe</a>	<a href="#">Executable in MS DOS window</a>
<a href="#">Run_g160app2.bat</a>	<a href="#">An example batch file for code evaluation with g160app2.exe</a>
<a href="#">Test</a>	<a href="#">Test vectors and reference results</a>

[The C-code has been compiled and tested with GNU gcc compiler in a Cygwin environment under Microsoft Windows. More detail information is available in the readme file in the electronic attachment.](#)

---