CELLULAR

SARA-G3 series

Audio

Application Note

Abstract

This document provides information and procedures to resolve audio related problems with SARA-G3 series modules.
In particular, it describes the procedures for tuning the hands-free algorithm (echo cancelation, automatic gain control, and noise reduction).

Document Information		
Title	**SARA-G3 series**	
Subtitle	Audio	
Document type	Application Note	
Document number	UBX-13001793	
Revision, date	R03	09-May-2016
Document status	Early Production Information	

Document status explanation	
Objective Specification	Document contains target values. Revised and supplementary data will be published later.
Advance Information	Document contains data based on early testing. Revised and supplementary data will be published later.
Early Production Information	Document contains data from product verification. Revised and supplementary data may be published later.
Production Information	Document contains the final product specification.

This document applies to the following products:

Product name
SARA-G340
SARA-G350

Contents

1 Introduction

This document provides information and procedures for the resolution of the potential audio related problems on SARA-G3 series modules. It also addresses the DTMF signaling decoder functionality available via the +UDTMFD AT command, implemented following the multi-part ETSI Standard ES 201 235 [6].

As an example, the document describes a procedure for tuning the hands-free algorithm (echo cancellation, Automatic Gain Control, noise reduction).

For a detailed description of audio parameters and AT commands, see the u-blox AT Commands Manual [1].

For a detailed description of SARA-G340 / SARA-G350 series module audio interface, see the SARA-G3 / SARA-U series System Integration Manual [2].

☞ An index finger points out key information pertaining to integration and performance.

⚠ **A warning symbol indicates actions that could negatively impact or damage the module.**

2 Introduction to HF (hands-free) algorithm tuning

After connecting external audio devices (i.e. microphone and loudspeaker) to the cellular module, the far-end user might hear an acoustic echo. This problem typically occurs when the device gain is set high to work at a distance (i.e. in hands-free application). SARA-G modules provide a hands-free algorithm to remove echo. The HF parameters control the algorithm within the uplink audio path in use (see the u-blox AT Commands Manual [1], +USPM and +UHFP AT commands), stored in the NVM dynamic parameters profile (see the section 3.1 and AT&W command in u-blox AT Commands Manual [1]).

Section 3 presents a step-by-step procedure for choosing parameters to remove echo heard on the far-end side. In case of HF systems with high echo coupling and high non-linearity on the loudspeaker (making the EC cancellation ineffective), a no duplex set-up is recommended, as presented in section 4.

2.1 HF algorithm description

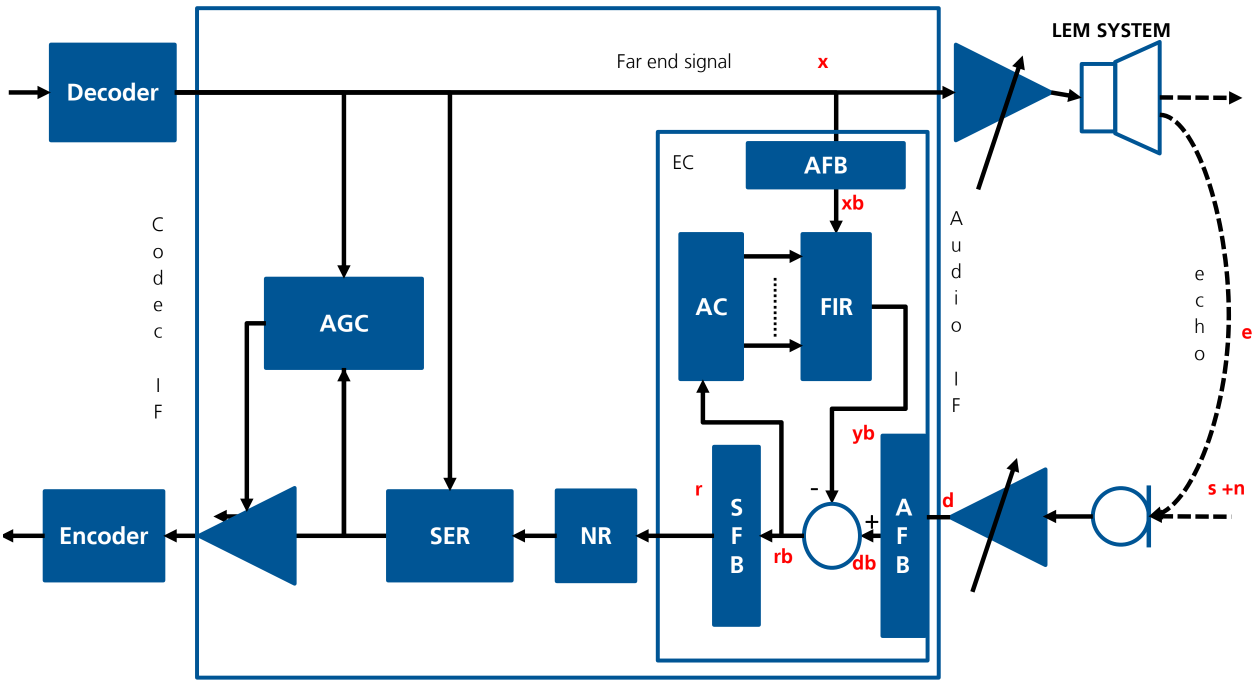

Figure 1: Block diagram for hands-free

The LEM system (Loudspeaker-Enclosure-Microphone) is a non-linear, time varying system. The microphone signal (d) is composed of near-end speech (s), echo (e) of the far end signal(x) and noise (n).

The echo canceller (EC) splits the downlink (x) and microphone (d) signals into sub-bands signals (xb, db) with 2 Analysis Filter Bank (AFB). For each sub-band a FIR filter emulates the LEM system behavior for the corresponding sub-band and generates an estimate (yb) of the acoustic echo produced by the signal (xb) on the LEM in that sub-band. For each sub-band the estimated echo (yb) is subtracted from the microphone sub-band signal (db).The residual echo (rb) of each sub-band is re-combined by a Synthesis Filter Bank (SFB) to a single full spectrum residual echo signal (r).

Since the LEM is time varying, the Adaptation Control block (AC) adapts the coefficients of each FIR filters; AC is a block NLMS adaptive algorithm based on the residual echo (rb) heard when the near-end speaker is silent (s=0). When the near-end speaker is not silent (double-talk condition) the filter adaptation is suspended. Full spectrum residual echo and noise (r) are then lowered by Noise Reduction (NR), Spectral Echo Reduction (SER) and Automatic Gain Control (AGC). AGC is disabled when in double-talk.

2.2 HF algorithm parameters

The HF parameters of the +UHFP AT command set the corresponding algorithm parameters. For the command description, see the u-blox AT Commands Manual [1].

2.2.1 Parameters for block activation and initialization

HF_ALGORITHM_INIT Range 0x0000, 0x3FFF

The audio driver uses this parameter to initialize the algorithm when a call starts. The parameter is a set of flags that control the activity and initialization of the EC, AGC and NR blocks. It can be set only in hexadecimal format.

Flag	Meaning
Bit #0 set	Echo cancellation (EC) initialization
Bit #1 set	EC restart (without coefficient initialization)
Bit #2 set	EC on
Bit #3 set	Noise reduction initialization
Bit #4 set	NR on
Bit #5 set	Dynamic Echo Suppression (DES) initialization. The DES module is disabled, therefore DE parameters have no influence.
Bit #6 set	DES on. The DES module is disabled, therefore DE parameters have no influence.
Bit #7 set	Automatic Gain Control (AGC) initialization
Bit #8 set	AGC on
Bit #9 set	Reconfigure
Bit #10 set	Unused
Bit #11 set	Unused
Bit #12 set	Spectral Echo Reduction (SER) Initialization
Bit #13 set	SER on

Table 1: HF_ALGORITHM_INIT flags explanation

Examples:

Configuration	Command	Remarks
EC only	`AT+UHFP=0,0x0005,,,,,,,0,0,500,8192,7500,7500,2,100,100,100,0,0,0`	
NR only	`AT+UHFP=0,0x0018,,,,,,,0,0,500,8192,7500,7500,2,100,100,100,0,0,0`	
AGC only	`AT+UHFP=0,0x0180,,,,,,,0,0,500,8192,7500,7500,2,100,100,100,0,0,0`	
EC+ SER	`AT+UHFP=0,0x3065,,,,,,,0,0,500,8192,7500,7500,2,100,100,100,0,0,0`	
EC+ AGC+SER	`At+UHFP=0,0x31E5,,,,,,,0,0,500,8192,7500,7500,2,100,100,100,0,0,0`	
EC+ NR + AGC + SER	`At+UHFP=0,0x31FD,,,,,,,0,0,602,8192,7500,7500,2,110,110,110,0,0,0`	
All off	`AT+UHFP=0,0x0000,,,,,,,0,0,500,8192,7500,7500`	Optional parameters can be omitted

Table 2: HF_ALGORITHM_INIT examples

2.2.2 Unavailable parameters

The parameters presented in this paragraph are not used. They are solely maintained in the command for backwards compatibility with LEON-G series modules.

In the information text response to the test command (AT+UHFP?) their value is shown as "NA" (Not Available).

In the set command, they can be omitted, e.g.:

AT+UHFP=<uplink_path_num>,<hf_algorithm_init>,,,,,,,<add_atten>,<min_atten>,<max_atten>,<nr_sw_2>,<nr_u_fak_0>, <nr_u_fak>

If they are not omitted, the value must be within the allowed range.

Parameter	Range	Applicability in SARA-G module
HF_ALGORITHM_RESTART	0x0000 to 0x07FF	Not Used
STEP_WIDTH	0 to 32767	Not Used
LMS_LENGTH	2 to 400	Not Used
LMS_OFFSET	0 to 400	Not Used
BLOCK_LENGTH	2, 4, 5 and 8	Not Used
RXTX_RELATION	-960 to 960	Not Used

Table 3: Not available parameters range

2.2.3 Parameters for automatic gain control (AGC)

Parameter	Range	Default value	Applicability in SARA-G module
ADD_ATTEN	-960 to 960	960	This value is added to the calculated attenuation as bias
MIN_ATTEN	0 to 960	0	Minimum attenuation of the microphone signal by the AGC. If calculated attenuation is lower than MIN_ATTEN, then attenuation is increased to MIN_ATTEN.
MAX_ATTEN	0 to 960	602	Maximum attenuation of the microphone signal by the AGC. If the calculated attenuation is higher than MAX_ATTEN, then the attenuation is decreased to MAX_ATTEN.

Table 4: AGC parameters description

The AGC parameters update the following attenuations:

- Additional Attenuation Level (dB) = 3/60* ADD_ATTEN
- Minimum Attenuation Level (dB) = 3/60* HF_MIN_ATTEN
- Maximum Attenuation Level (dB) = 3/60* HF_MAX_ATTEN

2.2.4 Parameters for noise reduction (NR)

Noise reduction operates on the eight frequency bands (band 0: 0-250 Hz; band 1: 250-750 Hz...band 7: 3250-3750 Hz). In band 0 the ear is less sensitive. For each band the NR computes a gain to apply (attenuation).

Parameter	Range	Default value	Remarks
NR_SW_2	0 to 32767	8192	This is the maximum attenuation that the NR can introduce. It is linear; where 32767 means 1 (0 dB; in this case no attenuation allowed, so there is no noise reduction). Very low values allow a strong attenuation but voice can result distorted (metallic). A good compromise is that the value is included in the range that goes from 4096 to 16384 (-18 to -6 dB).

Parameter	Range	Default value	Remarks
NR_U_FAK_0	0 to 16384	16384	This is the weighting factor for frequency band 0 (0 Hz-250 Hz). Increasing this factor will cause a better noise reduction in this band but also higher distortion of speech. Linear; weighting factor =NR_U_FAK_0 / 32768
NR_U_FAK	0 to 16384	16384	Factor of NR in the bands 1 to 7 (250 Hz -3750 Hz). This is the weighting factor for frequency band 0 (1 to 7 (250 Hz-3750 Hz)). Increasing this factor will cause a better noise reduction in this band but also higher speech distortion. Linear; weighting factor =NR_U_FAK_0 / 32768

Table 5: Noise reduction parameters description

Examples:

Configuration	Value	Remarks
NR_SW_2	4096	0.125 = -18 dB gain (18 dB is the maximum attenuation)
NR_U_FAK_0	16384	Weighting factor = 0.5
	8192	Weighting factor = 0.25
NR_U_FAK	16384	Weighting factor = 0.5
	4096	Weighting factor = 0.125

Table 6: Noise reduction parameters examples

2.2.5 Parameters for echo cancellation

Echo cancellation is a sub-band (SB) design, where identical systems on each sub-band perform adaptive linear filtering and subsequent echo subtraction (see section 2.1). After the echo subtraction, the sub-band signals are combined back to a single full spectrum signal. EC uses three sub-bands when the speech channel uses Narrow Band codec, six sub-bands with Wide Band codec.

Parameter	Range	Default value	Remarks
EC_BLOCK_LENGTH	1, 2, 4, 5, 8	4	LMS coefficient adaptation block length. It specifies the number of frames during which the adaptive filter coefficients are updated in the AC blocks. It can take only the values 1, 2, 4, 5 and 8 where 1 indicates updating of filter parameters at every frame (160 samples or 20 ms for narrow band), 4 represents updating every four frames. The higher this number, the slower but more accurate the adaptation converges.
EC_NR_COEFF_REAL	0 to 1100	220	Number of filter coefficients in the sub-band EC, for real sub band (in Narrow Band mode: 0 - 0.8 kHz in Wide Band mode: 0 - 0.73 kHz).
EC_NR_COEFF_COMPLEX1	0 to 1100	220	Number of filter coefficients in the sub-band EC, for complex sub band 1 (in Narrow Band mode: 0.8 - 2.4 kHz; in Wide Band mode: 0.73 - 2.18 kHz)
EC_NR_COEFF_COMPLEX2	0 to 1100	220	Number of filter coefficients in the sub-band EC, for complex sub band 2 (in Narrow Band mode: 2.4 - 4 kHz; in Wide Band mode: 2.18 - 3.64 kHz)
EC_NR_COEFF_COMPLEX3	0 to 1100	0	Number of filter coefficients in the sub-band EC, for complex sub band 3 (in Narrow Band mode: ignored; in Wide Band mode: 3.64 - 5.09 kHz)
EC_NR_COEFF_COMPLEX4	0 to 1100	0	Number of filter coefficients in the sub-band EC, for complex sub band 4 (in Narrow Band mode: ignored; in Wide Band mode: 5.09 - 6.56 kHz)
EC_NR_COEFF_COMPLEX5	0 to 1100	0	Number of filter coefficients in the sub-band EC, for complex sub band 5 (in Narrow Band mode: ignored; in Wide Band mode: 6.56 - 8 kHz)

Table 7: Echo cancellation parameters description

3 Procedure for echo canceller tuning

This is a step-by-step procedure to tune the audio path parameters for the removal of the echo heard on the far-end side.

See the u-blox AT Commands Manual [1] for more details on the AT commands and their parameters that are used in the tuning procedure (AT+USPM, AT+USGC, AT+UMGC, AT+UHFP, AT+UUBF, AT+UDBF, AT&W, AT&F, AT&Y). Check the path index for uplink and downlink to be used in these commands.

1 Turn off all the hands-free algorithm by means of the AT command:

AT+UHFP=<uplink_path_num>,0x0000,,,,,,,0,0,500,8192,7500,7500,1,100,100,100,60,60,60

2 In case of an hands-free device implementation, set the sidetone to 0 to avoid Larsen effect through the AT command:

AT+USTN=<downlink_path_num>,0

3 Regulate the gain on speaker and microphone used so that speech is not distorted on both uplink and downlink. This is very important because the echo cancellation algorithm only works efficiently in linear mode. Check (if possible also by oscilloscope) that the speech signal is not clipped. Tune the gain on the downlink path using AT+USGC command if the speech signal on speaker is distorted. Tune the gain on the uplink path using AT+UMGC command, if the speech signal from microphone is distorted.

4 Begin tuning the EC parameters with the AT command, while SER, AGC and NR are switched off.

AT+UHFP=<uplink_path_num>,0x0005,,,,,,,0,0,602,8192,16384,16384,4,100,100,100,0,0,0

Parameter	Value	Meaning
<hf_algorithm_init>	0x0005	Only echo cancellation initialization and on.
<ec_block_length>	4	Start updating adaptive filter coefficients every four frames. This is a good compromise between fast convergence and stability.
<ec_nr_coeff_real>	100	
<ec_nr_coeff_complex1>	100	
<ec_nr_coeff_complex2>	100	
<ec_nr_coeff_complex3>	0	
<ec_nr_coeff_complex4>	0	
<ec_nr_coeff_complex5>	0	

- Change one <ec_nr_coeff_*> parameter at a time and measure the resulting echo.
- Repeat the previous step to converge at a value that provides minimum echo. Perform these steps for all the <ec_nr_coeff_*> parameters.
- The <uplink_path_num> parameter corresponds to the index of the uplink path in use. Check the uplink path in use by command:

  AT+USPM?

  +USPM: <main_uplink>,<main_downlink>,<alert_sound>,<headset_indication>,<vmic_ctrl>

  Use <main_uplink> value for <uplink_path_num> in all the examples of this procedure.
- The EC parameters are optional. If omitted, the default values 4,220,220,220,0,0,0 are set.

  E.g: AT+UHFP=<uplink_path_num>,0x0005,,,,,,,0,0,500,8192,7500,7500

  sets the default EC parameters.
- Since AGC and NR are off (only echo cancellation initialized and on), any values in the allowed ranges of the AGC and NR parameters are accepted but not used; only the EC parameters are considered. Since the algorithm is adaptive, some seconds are needed to converge after that

AT+UHFP command is issued. After the algorithm convergence, a residual echo remains. Try to change the EC parameters till there is no residual echo. In very critical case, if the echo never disappears, try to find a minimum residual echo configuration.

- Use higher values of <ec_block_length> for more stable (but slower) convergence.

  E.g.: <ec_block_length>=1 convergence time is 1-2 s

  <ec_block_length>=4 convergence time is 4-6 s

  If NrCoeffs has been set properly to "small" values (<100) it should be a good compromise to set Blen=2

  If for any reason NrCoeffs was set higher than needed, Blen needs to be set to higher values than 2, and the following applies.
- Use higher values of <ec_nr_coeff_real> and <ec_nr_coeff_complex*> for a long reverberation time.
- Parameters <ec_nr_coeff_complex3>, <ec_nr_coeff_complex4>, and <ec_nr_coeff_complex5> are used only in WB speech. Test the EC performance both in NB and WB scenarios.
- If EC works correctly, a difference should be heard turning off the EC with the following AT command:

  AT+UHFP=<uplink_path_num>,0x0000,,,,,,,0,0,602,8192,16384,16384

5 Add the Dynamic Echo Suppression (DER) and Spectral echo Reduction (SER) algorithm to remove a residual echo, if present:

AT+UHFP=<uplink_path_num>,0x3065,,,,,,,0,0,602,8192,16384,16384,<ec_nr_coeff_real>,<ec_nr_coeff_complex1>,<ec_nr_coeff_complex2>,<ec_nr_coeff_complex3>,<ec_nr_coeff_complex4>,<ec_nr_coeff_complex5>

Parameter	Value	Meaning
<hf_algorithm_init>	0x3065	Echo cancellation initialization and on; SER initialization and on; DER initialization and on.
<ec_nr_coeff_real>		As in step 4.
<ec_nr_coeff_complex*>		As in step 4.

6 Add the AGC algorithm to remove a minimal residual echo, if present:

AT+UHFP=<uplink_path_num>,0x31E5,,,,,,,0,0,600,8192,16384,16384,<ec_nr_coeff_real>,<ec_nr_coeff_complex1>,<ec_nr_coeff_complex2>,<ec_nr_coeff_complex3>,<ec_nr_coeff_complex4>,<ec_nr_coeff_complex5>

Parameter	Value	Meaning
<hf_algorithm_init>	0x31E5	Echo cancellation initialization and on; AGC initialization and on; SER initialization and on; DES initialization and on.
<add_atten>	0	0 dB minimum AGC attenuation.
<min_atten>	0	0 dB additional AGC attenuation.
<max_atten>	600	600*3/60 = 30 dB maximum AGC attenuation.
<ec_nr_coeff_real>		As found in step 4.
<ec_nr_coeff_complex1>		As found in step 4.
<ec_nr_coeff_complex2>		As found in step 4.
<ec_nr_coeff_complex3>		As found in step 4.
<ec_nr_coeff_complex4>		As found in step 4.
<ec_nr_coeff_complex5>		As found in step 4.

7 If the residual echo is still present, try to use higher <min_atten> values:

AT+UHFP=<uplink_path_num>,0x31E5,,,,,,,50,100,500,8192,16384,16384,<ec_nr_coeff_real>,<ec_nr_coeff_complex1>,<ec_nr_coeff_complex2>,<ec_nr_coeff_complex3>,<ec_nr_coeff_complex4>,<ec_nr_coeff_complex5>

Parameter	Value	Meaning
<add_atten>	50	50*3/60 = 2.5 dB additional AGC attenuation
<min_atten>	100	100*3/60 = 5 dB minimum AGC attenuation
<max_atten>	500	500*3/60 = 25 dB maximum AGC attenuation

Example of AGC settings:

AGC	<add_atten>	<min_atten>	<max_atten>
Weak AGC	0	0	200
Moderate AGC	100	100	500
Strong AGC	200	200	500
No-duplex AGC	500	500	500

⚠ **Using strong AGC setting can decrease the performance in double-talk scenario and lead to a no-duplex configuration (see section 4).**

8 Add the NR algorithm to remove a residual noise on the uplink path (if present):

AT+UHFP=<uplink_path_num>,0x31FD,,,,,,,<add_atten>,<min_atten>,<max_atten>,8192,7500,7500,<ec_nr_coeff_real>,<ec_nr_coeff_complex1>,<ec_nr_coeff_complex2>,<ec_nr_coeff_complex3>,<ec_nr_coeff_complex4>,<ec_nr_coeff_complex5>

Parameter	Value	Meaning
EC parameters		As found in step 4, 5 and 6.
AGC parameters		
<hf_algorithm_init>	0x31FD	Echo cancellation initialization and on; NR initialization and on; AGC initialization and on; SER initialization and on; DES initialization and on.
<nr_sw_2>	4096	20 log (4096/32767) = -18 dB minimum NR attenuation
<nr_u_fak_0>	7500	7500/32768 = 0.23 weighting factor for frequency band 0
<nr_u_fak>	7500	7500/32768 = 0.23 weighting factor for frequency band 1-7

9 As with the previous step, the NR parameters can be changed until the minimum residual noise remains on the uplink path. To appreciate the NR effect, listen to the uplink speech in silence scenario (neither uplink nor downlink speech, nor echo present) and switch only the NR off/on by these commands:

AT+UHFP=<uplink_path_num>,0x0000,,,,,,,0,0,500,<nr_sw_2>,<nr_u_fak_0>,<nr_u_fak>
AT+UHFP=<uplink_path_num>,0x0018,,,,,,,0,0,500,<nr_sw_2>,<nr_u_fak_0>,<nr_u_fak>

- Rule to get proper quality during speech periods:
  <nr_u_fak_0> **<** 2* <nr_sw_2>
  <nr_u_fak> **<** 2* <nr_sw_2>
- Burst of noise during speech periods when:
  <nr_u_fak_0> **>** 2* <nr_sw_2>
  <nr_u_fak> **>** 2* <nr_sw_2>
- Metallic voice if:
  <nr_u_fak_0> **<** <nr_sw_2>
  <nr_u_fak> **<** <nr_sw_2>

3.1 Storing the parameters in the profile

☞ Follow this procedure to save the parameters in the NVM dynamic parameters:

1. Write the run-time configuration to NVM with the AT&W command (e.g. AT&W0; more details in the u-blox AT Commands Manual [1]).
2. Assure the boot loading is performed with the desired parameter profile (e.g. profile 0 if the parameter saving was performed with AT&W0; use AT&Y0 to select this).
3. Save the run-time configuration to Flash memory with the AT+CPWROFF command.
4. Switch on of the device.

4 No-duplex configuration

The no-duplex setup is recommended for HF systems with high echo coupling and high non-linearity on loudspeaker which make the EC cancellation ineffective.

☞ The no-duplex configuration particularly makes use of the AGC only as muting/un-muting device of the TX path.

The operating conditions of this no-duplex configuration are:

- Far-end user speaking: a high and constant attenuation is added to Tx signal; Tx signal is almost muted.
- Far-end user silent: Tx signal is un-muted (AGC off), both when near-end user is speaking or silent.

4.1 Procedure

Configure the AGC to a high attenuation performance, e.g.:

AT+UHFP=<uplink_path_num>,0,0x0180,,,,,,500,500,500,8192,7500,7500		
Parameter	**Value**	**Meaning**
<hf_algorithm_init>	0x0180	Only AGC initialization and on.
<add_atten>	500	500*3/60 = 28 dB additional AGC attenuation
<min_atten>	500	500*3/60 = 28 dB minimum AGC attenuation
<max_atten>	500	500*3/60 = 28 dB maximum AGC attenuation

5 DTMF decoder

5.1 About ETSI DTMF

The dual-tone multi-frequency (DTMF), also known as Touch Tone, is used for telephone signaling over the line in the voice frequency band to the local exchange. Prior to DTMF the phone systems had used a system known as pulse dialing to dial numbers, which works by rapidly disconnecting and connecting the calling party's phone line, like turning a light switch on and off.

The multi-part ETSI Standard ES 201 235 [6] specifies how to apply DTMF signaling to transmitters and receivers. It conforms to the International Telecommunication Union (ITU-T) Recommendation Q.23 [3] and it provides a complete set of requirements for all the applications intending to use DTMF signaling.

The level of detail enables manufacturers of telecommunications equipment incorporating DTMF signaling to design the equipment such that it facilitates highly reliable signaling. It applies to the DTMF signaling in the local access network, in which the transmission path between transmitter and receiver corresponds to a 2-wire analogue subscriber line, as well as to DTMF signaling over an end-to-end transmission path in the telecommunication network.

5.2 About DTMF

The dual-tone multi-frequency signaling is a standard in telecommunication systems. It has been gaining popularity for some years now because of its numerous advantages over the traditional telephone signaling scheme. In the DTMF scheme, a telephone is equipped with a keypad as shown in Figure 2. The A, B, C, and D keys are usually not present on a regular telephone keypad. Each key represents the sum of a pair of tones. One tone is from the high-frequency group between 1 kHz and 2 kHz, and the other tone is from the low-frequency group below 1 kHz. These frequencies are selected carefully so that the DTMF signal, which is the sum of the two tones, can be clearly distinguished as the signaling tone even in the presence of speech waveforms that might occur on the line.

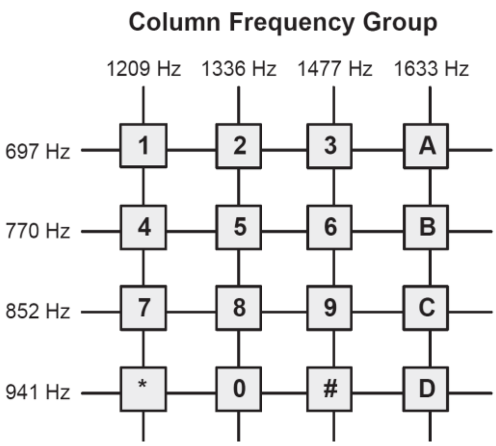

Figure 2: Touch-Tone telephone keypad: a row and a column tone is associated with each digit

5.2.1 The DTMF signal definitions

The tone frequencies, as defined by the Precise Tone Plan, are selected such that harmonics and inter-modulation products do not cause an unreliable signal. The frequency is not a multiple of another, the difference between any two frequencies does not equal any of the frequencies, and the sum of any two frequencies does not equal any of the frequencies. The frequencies were initially designed with a ratio of 21/19, which is slightly less than a whole tone.

6 DTMF signaling decoder on cellular modules

The u-blox cellular modules can be configured to perform DTMF detection on the RX speech channel. The DTMF decoder is part of the In-Band modem feature and the +UDTMFD AT command is used to configure it (for more details on the command description, see the u-blox commands manual [1]).

6.1 Implementation

Enable the DTMF decoder via the AT command once per module power cycle and before any call set up, e.g.:

```
AT+UDTMFD=1,2
OK
```

At each call setup, the DTMF decoder is automatically enabled. During the call, the DTMF decoder provides URCs for each detected digit, e.g.:

```
+UUDTMFD: 4
```

Here, the digit "4" has been detected.

6.2 Performance criteria

Various standards bodies (ITU-T, ETSI, EIA/TIA), mobile network operators (NTT, AT&T) and other players in the communication industry (MITEL, Bellcore) have established different performance tests and criteria for DTMF decoders.

The u-blox decoder implementation has followed the ETSI specifications as described in the multi-part ETSI Standard ES 201 235 [6]. However, the AT interface allows the decoder configuration for the performance criteria customization on need.

There are two main performance indicators for DTMF detectors:

- **Detection performance** – is the ability to correctly decode the DTMF tones in various network conditions. The modern networks use compression which introduces distortions that may invalidate at detector input a correctly generated DTMF tone.
- **Speech immunity** – is the DTMF talk-off abatement performance. Talk-off is the term that describes when a human voice is able to trigger DTMF tones during a telephone call. Talk-off occurs when the DTMF detector tries to translate sounds into DTMF tones causing false detections.

The decoder performance is also characterized by the robustness towards digit repetitions (special case of false detection), for instance those caused by interruptions in the DTMF tones. The ETSI standard specifies that a detected digit shall be unaffected by disturbances having a duration of less than 20 ms. Nevertheless, such a criterion can be not sufficient to avoid false digit repetitions in case of networks characterized by high distortions or speech frame losses.

In some conditions, the overall performance may be improved by increasing the tone duration and the pause between tones (inter-digit interval); in this way, the performance is higher if there is the possibility to decrease the digit transmission rate and tune the detector accordingly.

In general, the higher the speech immunity, the higher the risk of missed detections. The right trade-off between detection performance and talk-off abatement performance depends on the application.

6.2.1 Decoder configuration

At each module power cycle, the decoder is configured with factory-programmed values.

The AT command can reconfigure the values at any time, even run-time during a call, e.g.:

```
AT+UDTMFD=1,2,4,400,10,3
```

The decoder has six configuration parameters:

Parameter	**Range**	**Default value**	**Description**
<urc_en>	0: disable 1: enable	N.A.	Enables the URCs on a specific AT terminal. Mandatory parameter
<mode>	0: disabled 1: normal 2: robust	N.A.	DTMF feature enabling/disabling and activation mode definition.
<att_cfg>	0-15	4	Controls the accepted signal levels. The signal is scaled down by 24 dB at the detector input.
<threshold>	100-10000	400	Controls the accepted signal levels. The digit recognition starts when the output of the analysis filter bank reaches the value of 400
<immunity>	0-20	14	Calibrates the speech immunity strength.
<max_int>	1-255	2	Controls the false digit repetitions. The expected minimum pause between the digits is 40 ms. Maximum signal interruption is 20 ms.

Table 8: +UDTMD parameter description and factory-programmed values

The factory-programmed values may vary in different products or product versions.

☞ By default, it is suggested to activate the DTMF in robust mode.

☞ <att_cfg> and <threshold> default values are optimized for the best performance in terms of signal level operating range, and complying with the ETSI requirements.

☞ On SARA-G3 "00S", "00X", "01S" and "01B" product versions the <mode> parameter is mandatory.

6.2.1.1 Activation mode (<mode> parameter)

The detector can be activated in the normal and the so-called robust mode. The robust mode is characterized by a reduced risk of any kind of false detections. The robustness is achieved by analyzing the input signal in the time domain. In fact human voice, melodies and other signals, as well as speech codecs like AMR or other disturbances from the network that potentially cause false detections, are generally "touchtone-like generators" for very short time.

The DTMF detector in robust mode meets the ETSI expectations in terms of detection performance ETSI TR 126 975 [7].

Robust mode advantages: case study

Figure 3 shows a normalized signal as presented at the DTMF decoder input, corresponding to the digit sequence "1,2,3,4,5". The signal has been generated by key clicks (approx. 500 ms) on a VoIP telephone connected with a u-blox cellular module. The short burst before every tone is actually the start of the tone itself that is interrupted by the network after ~ 25/30 ms and restored after ~ 100 ms.

The false digit repetitions affect the normal mode. The short bursts are detected as an independent DTMF tone, thus the detector output is "1,1,2,2,3,3,4,4,5,5". In robust mode, the detector rejects the bursts and correctly outputs the digits.

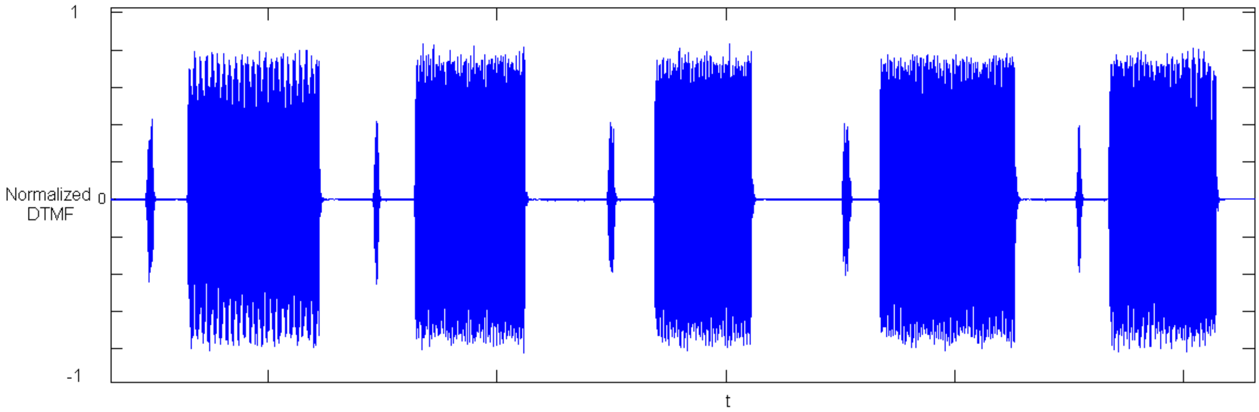

Figure 3: Normalized tones corresponding to digit presses "1, 2, 3, 4, 5", with interruptions caused by VoIP-based connection

6.2.1.2 Accepted signal levels (<att_cfg> and <threshold> parameter)

The **<att_cfg>** parameter applies attenuation to the input signal in steps of 6 dB: 0 for 0 dB attenuation, 15 for 90 dB attenuation (which corresponds to mute signal at decoder input). In general, the parameter can be used to adapt the decoder to special network conditions (e.g. extremely high or extremely low tone levels). <att_cfg> should be configured as low as possible but avoiding overflows. There is an overflow protection mechanism that automatically scales down signals that lead the detector into algorithmic overflow. The automatic scaling is acknowledged through the following URC:

```
+UUDTMFDE: 1
```

The new attenuation can be retrieved getting the last parameter of the read command, e.g.

```
AT+UDTMFD?
+UDTMFD: 1,2,4,100,14,2,5
OK
```

Here, the attenuation has been increased from 4 to 5. The overflow protection mechanism only increases the attenuation. If overflow is notified, it is not guaranteed that the decoder performs in best condition and additional attenuation might be required. A detection result prior to an overflow notification is considered unreliable.

The **<threshold>** parameter is the current threshold applied to the signal level in order to be considered valid (i.e. to enter the operating condition). Higher thresholds give better performance especially in terms of speech immunity and false detections, with the cost of the increase of the lower boundary of the operating range. This parameter is not expressed in dB.

There is a relationship between <att_cfg> and <threshold>: the decoder performance at a specific signal level does not change if for each 6 dB attenuation increase the threshold is doubled (which corresponds to a 6 dB increase, too).

In general, it should not be necessary to change the <att_cfg> and <threshold> parameters. Nevertheless, standards and operators may require slightly different operating conditions for DTMF. Generally, ranges of 20-30 dB are required, while the not-operating condition may vary from -29 to -40 dBm.

The default values are selected to meet operating conditions according to the requirements from ETSI ES 201 235-4 [5], for more details see the Table 9.

Conditions with default values (in robust mode)	Valid	Not Valid
Signal level x (dBm0)	-36 <= x <= -3	x < -40

Table 9: Default values handling

The levels are expressed in decibels with respect to 0x7FFF clipping value.

6.2.1.3 Immunity/talk-off abatement (<immunity> parameter)

The **<immunity>** parameter calibrates the decoder with respect to speech immunity performance: 0 for no immunity, 20 for maximum immunity performance. Unlike the robust mode, the talk-off abatement algorithm is based on the spectral analysis of the signal.

For certain end-to-end applications in which the talk-off abatement is not relevant (since voice or other disturbing signals are not injected in the voice channel), the speech immunity can be lowered or even completely disabled, having the advantage of an improved detection performance, for instance with low bit-rate codecs.

The default immunity value (14), combined with the robust mode, complies with the ETSI requirements for speech immunity ETSI ES 201 235-4 [5]. According to ETSI ES 201 235-3 [4], ETSI ES 201 235-4 [5], "Table 2: Signal condition requirements, NOTE2", the talk-off performance is not directly specified as set of requirements for the existence or non-existence of signal conditions. The performance is indirectly specified through the speech immunity requirements of clause 4.2 in ETSI ES 201 235-3 [4] and ETSI ES 201 235-4 [5], "Speech immunity performance".

Table 10 provides the tests results with four different detector configurations with respect to speech immunity (default values for the other parameters used). Talk-offs represent the number of false detections during the testing.

<immunity> parameter value	<mode> parameter value	Talk-offs
0	normal mode	9900
0	robust mode	4800
14	normal mode	100
14	robust mode	5 (as in ETSI reference)

Table 10: Test results with four different detector configurations

6.2.1.4 False digit repetitions (<max_int> parameter)

The network conditions can generate more or less short interruptions of tones that may cause false detections – digit repetitions. ETSI requires that a decoder is unaffected by disturbances a duration of less than 20 ms, which may be not sufficient for network conditions.

The **<max_int>** parameter allows the tuning of the maximum interruption that a detected tone may have, such that is still interpreted as a single digit and thus avoiding false digit repetitions.

The <max_int> parameter also represents the minimum expected pause between two DTMF tones. Therefore if a decoder is configured to compensate interruptions up to e.g. 80 ms (<max_int>=4), the DTMF transmitter shall be configured to generate tones with a pause between them larger than 80 ms, otherwise the decoder recognizes two subsequent tones associated to the same digit as a single digit.

By default <max_int> is set to 40 ms, which is ETSI compliant according to receiver's digit recognition condition requirement in ETSI ES 201 235-4 [5], cit. "any tone shall be preceded by the continuous absence of a valid signal condition for more than 40 ms".

6.2.1.5 Not configurable signal condition and tolerances / default values

Table 11 reports the not configurable signal conditions. The signal conditions and tolerances comply with ETSI ES 201 235-3 [4] and ETSI ES 201 235-4 [5].

Signal conditions an tolerances	Valid	Not Valid
Frequency Deviation	<= ± (1,5% + 2) Hz	
Twist (signal level difference)	< 12 dB	
Reverse Twist (signal level difference)	< 12 dB	

Table 11: Not configurable factory-programmed signal conditions and tolerances on SARA-G modules

- Twist: the lower tones are higher in amplitude than the higher tones
- Reverse twist: the lower tones are lower in amplitude than the higher tones

6.3 DTMF performance measurements

DTMF performance is measured with respect to speech immunity and detection performance.

6.3.1 Speech immunity

Speech immunity tests have been performed according to ETSI ES 201 235-4 [5] (Paragraph 4.3, Annex A and Annex B), connecting the u-blox module with a network simulator using a full-rate speech codec. The test results, presented in the Table 12, have been obtained with factory-programmed configuration values, only varying the <mode> and <immunity> parameters. The full speech immunity is reached if the DTMF detector has maximum 5 talk-offs (i.e. false detections caused by 20 minutes of speech-like test signal injected into the detector).

<immunity> parameter value	<mode> parameter value	Talk-offs
0	normal mode	9900
0	robust mode	4800
14	normal mode	100
14	robust mode	5 (as in ETSI reference)

Table 12: ETSI Speech immunity tests, with 20 minutes test signal

☞ The factory-configured DTMF decoder activated in robust mode passes the speech immunity test.

6.3.2 Detection performance

The detection performance measurement and benchmarking was done as in ETSI TR 126 975 [7], Chapter 10 "Performances with DTMF tones", the tests implemented as in the described test procedure, on a sub-set of experiments. This ETSI document is not intended to be a DTMF decoder specification. Rather, it evaluates the transparency of the FR and AMR speech codecs to DTMF tones.

The benchmarking with the ETSI reference DTMF decoder is considered a valid performance measurement. It points out the problems that necessitate the widely used speech codecs adopted by 2G and 3G cellular networks: **the AMR low bit rate modes are not transparent to DTMF tones** (see the ETSI TR 126 975 [7] and Table 14).

The tests have been performed with factory-programmed configuration values, only varying the <mode> and <immunity> parameter. Five different experiments from ETSI TR 126 975 [7] at various signals levels and with or without frequency deviation and reverse twist have been considered. Each experiment is made up of 20 repetitions of a sequence of 16 DTMF digits with tone 80 ms duration and 80 ms pause duration.

Table 13 and Table 14 illustrate the DTMF decoder performance with respect to two different speech codecs:

- FR GSM 13 kb/s codec
- AMR 4.75 kb/s codec

6.3.2.1 Full rate GSM 13 kb/s codec

Each element in Table 13 reports the percentage of undetected digits and the percentage of false detections. For each x/y table element, x represents the percentage of undetected DTMF digits and y represents the percentage of out-of-sequence digits (false detections).

FR GSM 13 kb/s	<immunity>=0 normal mode	<immunity>=0 robust mode	<immunity>=14 normal mode	<immunity>=14 robust mode	ETSI reference
exp7: -6 dBm	0/30	0/0	0/0	0/0	**0/0**
exp8: -16 dBm	0/22	0/0	0/0	0/0	**0/0**
exp9: -26 dBm	0/8	0/0	0/0	0/0	**0/0**
exp10: -16 dBm+frequency deviation	0/8	0/0	0/0	0/0	**0/0**
exp11: -13 dBm with -6 dB (reverse) twist	0/16	0/0	0/0	0/0	**0/0**

Table 13: Results for each experiment (rows) for each decoder configuration (columns)

☞ 100% of detections are achieved with the factory-programmed detector in both normal and robust mode, without false detections.

False detections are present only in normal mode with completely disabled immunity (<immunity>=0). This configuration, which represents a configuration at boundary conditions, is not recommended.

6.3.2.2 AMR 4.75 kb/s codec

Each element in Table 14 reports the percentage of undetected digits and the percentage of false detections. For each x/y table element, x represents the percentage of undetected DTMF digits and y represents the percentage of out-of-sequence digits (false detections).

AMR 4.75 kb/s codec	<immunity>=0 normal mode	<immunity>=0 robust mode	<immunity>=14 normal mode	<immunity>=14 robust mode	ETSI reference
exp7: -6 dBm	0/24.0	0/0	4.7/0	20.9/0	**21.3/0**
exp8: -16 dBm	0/7.8	0/0	1.6/0	22.1/0	**24.8/0**
exp9: -26 dBm	0/0.6	0/0	1.8/0	19.7/0	**27.5/0**
exp10: -16 dBm+frequency deviation	0/12.5	0.3/0	1.6/0	19.0/0	**26.9/0**
exp11: -13 dBm with -6 dB (reverse) twist	0/8.1	16.6/0	11.8/0	34.7/0	**35.9/0**

Table 14: Results for each experiment (rows) for each decoder configuration (columns)

☞ Better detection performance than the ETSI reference is achieved with the factory-programmed detector in robust mode (23.28% vs 27.28% in average).

☞ Almost 100% of detections are achieved if the immunity in robust mode is disabled.

☞ exp11 with artificially added negative twist represents unreal/rare network situations. For more details, see the section 6.3.2.3.

6.3.2.3 Discussion

ETSI-compliant detector

The DTMF decoder in robust mode and default setting (<immunity> = 14) performs as expected by ETSI requirements, both with respect to speech immunity and detection performance.

AMR transparency

ETSI verified that low-bit rate codecs, in particular the AMR 4.75 kb/s codec, are not transparent to DTMF tones, especially the shorter ones, if an ETSI-compliant decoder is used. For instance, the AMR codecs have a tendency to add negative twist to DTMF signals. This is revealed by results of experiment exp11, in which an additional negative twist of 6 dB has been artificially added to DTMF tones prior AMR encoding. The DTMF factory setting for twist valid condition has been relaxed from the minimum recommended of 6 dB in ETSI ES 201 235-4 [5] to 12 dB: for more details see the section 6.2.1.5. Nevertheless, the exp11 signal conditions can be considered really boundary conditions, which are rare in real network situations.

Immunity configuration

With a cost of a reduced speech immunity performance, the u-blox DTMF decoder can be tuned to be more or less transparent to speech codec modes, acting on the <immunity> parameter. In particular, with the disabled immunity (<immunity>=0), it can cope with distortions introduced by the AMR 4.75 kb/s codec maintaining a detection performance close to 100%. The reduced immunity performance can be acceptable in controlled conditions of talk-off sources. A typical application which does not need speech immunity performance is the terminal end-to-end signaling, in which the microphone at DTMF generator side is disabled.

Normal mode

The normal mode combined with a proper level of immunity can give the right balance between the detection performance and the speech immunity performance. See for example the detection results with AMR codec, <immunity>=14 in normal mode detection, close to 98% hits without false detections, 100 talk-offs.

Tone duration

It is a recommendation for the transmitter. For end-to-end signaling, especially with low-bit-rate codecs, a minimum of 80 ms for tone duration is recommended. There are generally no benefits in having tones lasting more than 120 ms (on the contrary, the risk of false digit repetitions is increased).

Pause duration and <max_int>

It is a recommendation for the transmitter. ETSI recommends that if the transmitter automatically controls the DTMF signaling pause duration, the duration of the pause between any individual DTMF tone combination shall not be less than 65 ms. On need, the <max_int> can be configured accordingly to the transmitter's configuration, as proposed in the section 6.2.1.4.

Half-Rate (HR) codecs

The half-rate codecs may dramatically worsen the decoder performance. As stated by ETR 229 [8], "a serious commercial application using DTMF in the speech channel should not be supported with the GSM half rate codec.". This statement is valid for any codec working on half-rate channels, like for instance the HR-AMR (Half-Rate AMR). The half-rate speech channels are not only characterized by the distortions of low-bit rate codecs, but also by a higher error rate since the actual payload data rate is halved with respect to the full rate channel (for example, 6.5 kb/s vs 13kb/s).

☞ The module can be configured to not perform calls on half-rate channels through AT+UDCONF=30 command (see the section 6.4).

6.4 Configuration examples

The performance estimates of the following configuration examples are given for error free conditions (no speech frame drops). The frame drops may cause false digit repetitions that can be coped with <max_int> parameter configuration as discussed in the section 6.2.1.4.

6.4.1 ETSI-compliant decoder

It is achieved by the decoder enabled in robust mode with factory-programmed parameters:

```
AT+UDTMFD=1,2
OK
AT+UDTMFD?
+UDTMFD: 1,2,4,400,14,2,4
OK
```

Characterized by full speech immunity, the detection rate of this configuration can be less than 100% with low bit-rate codecs, as presented in the section 6.3.

6.4.2 ETSI-compliant decoder with guaranteed speech channel QoS

To get rid of low-bit rate codecs distortions, it is possible to configure the module to support and make calls only with a reduced speech codec set. The +UDCONF=30 AT command configures the speech codec. For more details, see the u-blox AT commands manual [1].

6.4.2.1 EFR, FR codec set restriction

For example, the ETSI-compliant decoder working with Enhanced Full Rate (EFR) and Full Rate (FR) codecs only guarantees a 100% detection performance with full speech immunity:

```
AT+UDCONF=30,6
OK
AT+UDTMFD=1,2
OK
```

6.4.2.2 Full-rate channel restriction

It is possible to restrict the channels only excluding half-rate channels; e.g.

```
AT+UDCONF=30,7
OK
```

This configures the module to use FR, EFR and FR-AMR codecs.

☞ For the description of +UDCONF=30 command for speech codec configuration, see the u-blox AT commands manual [1].

6.4.3 Custom DTMF detectors for low quality speech channels

A DTMF decoder can be configured to provide good performance even with low bit-rate codecs, at a cost of lower speech immunity or restricted operating range. In both cases the transmitter shall work in controlled condition.

The following configurations are guidelines and need actual in-field tuning and validation.

6.4.3.1 Reduced speech immunity

A good detection performance with low bit-rate codecs can be reached just by turning on the decoder in normal mode and keeping the factory-programmed parameters:

```
AT+UDTMFD=1,1
OK
AT+UDTMFD?
+UDTMFD: 1,1,4,400,14,2,4
OK
```

According to the performance measurements, the talk-offs increase statistically from 5 to 100, while the detection rate increases from approx 75% to 98% with the worst AMR codec case (4.75 kb/s). Varying the immunity parameter, the balance between talk-offs and detection rate can be differently distributed.

If the talk-off performance is not an issue, the immunity can be completely disabled and robust mode turned on (to avoid false detections).

```
AT+UDTMFD=1,2,,,0,2
OK
AT+UDTMFD?
+UDTMFD: 1,1,4,400,0,2,4
OK
```

The detection rate on worst AMR case should now be close to 100%.

6.4.3.2 Reduced operating range

The operating range reduction improves the speech immunity and in general performance vs false detection, e.g.:

```
AT+UDTMFD=1,1,,1200
OK
AT+UDTMFD?
+UDTMFD: 1,1,4,3200,14,2,4
OK
```

This is a configuration that should have an improved talk-off with respect to section 6.4.3.1 at the cost of reduction of a minimum operating level shift from approx -36 dBm to -18 dBm.

Appendix

A List of Acronyms

Abbreviation / Term	Explanation / Definition
AC	Adaptation Control block
AFB	Analysis Filter Bank
AGC	Automatic Gain Control
AP	Application Processor
AT	AT Command Interpreter Software Subsystem, or attention
DBF	Downlink Biquad Filters command
DES	Dynamic Echo Suppressor
DSP	Digital Signal Processing
EC	Echo cancellation algorithm
FIR	Finite Impulse Response filter
HF	Hands free Algorithm
LEM	Loudspeaker-Enclosure-Microphone
LMS	Least Mean Square
MGC	Microphone Gain control command
NLMS	Normalized Least Mean Square
NR	Noise Reduction algorithm
NVM	Non Volatile Memory
PSTN	Public Switched Telephone Network
RX	Receiver
SGC	Speaker Gain control command
SER	Spectral Echo Reduction
SFB	Synthesis Filter Bank
TX	Transmitter
UBF	Uplink Biquad Filters command

Related documents

[1] u-blox AT Commands Manual, Docu No UBX-13002752

[2] SARA-G3 series System Integration Manual, Docu No UBX-13000995

[3] ITU-T Recommendation Q.23: Technical features of push-button telephone sets

[4] ETSI ES 201 235-3 V1.3.1, Specification of Dual Tone Multi-Frequency (DTMF) Transmitters and Receivers; Part 3: Receivers

[5] ETSI ES 201 235-4 V1.3.1, Specification of Dual Tone Multi-Frequency (DTMF) Transmitters and Receivers; Part 4: Receivers for use in Terminal Equipment for end-to-end signaling

[6] Work Items with ETSI Document Number of "201 235", see Work Program search database, http://www.etsi.org/

[7] ETSI TR 126 975 V10.0.0 (2011-04), Performance characterization of the Adaptive Multi-Rate (AMR) speech codec (also 3GPP TR 26.975 version 10.0.0 Release 10)

[8] ETR 229: October 1995 (GSM 06.08 version 4.0.0), Performance characterization of the GSM half rate speech codec

[1] and [2] are available on our homepage (http://www.u-blox.com).

☞ For regular updates to u-blox documentation and to receive product change notifications, register on our homepage.

Revision history

Revision	Date	Name	Status / Comments
R01	02-Oct-2012	gbat	Initial release Last revision with document number GSM.G2-CS-13001
R02	06-Jun-2014	smos	Style and CI changes only
R03	09-May-2016	ague	Improved description of echo canceller tuning

Contact

For complete contact information visit us at www.u-blox.com

u-blox Offices

North, Central and South America

u-blox America, Inc.

Phone: +1 703 483 3180
E-mail: info_us@u-blox.com

Regional Office West Coast:

Phone: +1 408 573 3640
E-mail: info_us@u-blox.com

Technical Support:

Phone: +1 703 483 3185
E-mail: support_us@u-blox.com

Headquarters
Europe, Middle East, Africa

u-blox AG

Phone: +41 44 722 74 44
E-mail: info@u-blox.com
Support: support@u-blox.com

Asia, Australia, Pacific

u-blox Singapore Pte. Ltd.

Phone: +65 6734 3811
E-mail: info_ap@u-blox.com
Support: support_ap@u-blox.com

Regional Office Australia:

Phone: +61 2 8448 2016
E-mail: info_anz@u-blox.com
Support: support_ap@u-blox.com

Regional Office China (Beijing):

Phone: +86 10 68 133 545
E-mail: info_cn@u-blox.com
Support: support_cn@u-blox.com

Regional Office China (Chongqing):

Phone: +86 23 6815 1588
E-mail: info_cn@u-blox.com
Support: support_cn@u-blox.com

Regional Office China (Shanghai):

Phone: +86 21 6090 4832
E-mail: info_cn@u-blox.com
Support: support_cn@u-blox.com

Regional Office China (Shenzhen):

Phone: +86 755 8627 1083
E-mail: info_cn@u-blox.com
Support: support_cn@u-blox.com

Regional Office India:

Phone: +91 80 4050 9200
E-mail: info_in@u-blox.com
Support: support_in@u-blox.com

Regional Office Japan (Osaka):

Phone: +81 6 6941 3660
E-mail: info_jp@u-blox.com
Support: support_jp@u-blox.com

Regional Office Japan (Tokyo):

Phone: +81 3 5775 3850
E-mail: info_jp@u-blox.com
Support: support_jp@u-blox.com

Regional Office Korea:

Phone: +82 2 542 0861
E-mail: info_kr@u-blox.com
Support: support_kr@u-blox.com

Regional Office Taiwan:

Phone: +886 2 2657 1090
E-mail: info_tw@u-blox.com
Support: support_tw@u-blox.com