

# AN11855

## Voice Detection and Recognition

2.0 — 30 September 2016

Application note

### Document information

Info	Content
<b>Keywords</b>	LPC5411x, Voice Detection, Voice Recognition
<b>Abstract</b>	The application note explains the advantages of the LPC5411x ARM Cortex-M4F microcontroller family for a product which requires voice detection and voice recognition features.



## Revision history

Rev	Date	Description
2.0	30 September 2016	Revised, More technical details, NDA maybe required
1.0	21 June 2016	First release

## Reference Documents

Ref	Source	Title
[1]	NXP Semiconductors	LPC5411x data sheet
[2]	NXP Semiconductors	LPC5411x user manual
[3]	NXP Semiconductors	OM13090 user manual
[4]	ARM Ltd.	Cortex-M4_ReferenceManual
[5]	ARM Ltd.	Cortex-M4 Technical Reference Manual Revision r0p1

## Contact information

For additional information, please visit: <http://www.nxp.com>

For sales office addresses, please send an email to: [salesaddresses@nxp.com](mailto:salesaddresses@nxp.com)

## 1. Introduction

---

The LPC5411x are ARM Cortex-M4F based microcontrollers for embedded applications. These devices include an optional ARM Cortex-M0+ coprocessor, up to 192 KB of on-chip SRAM and up to 256 KB on-chip flash. Among other features, the LPC5411x family includes a Full Speed USB device interface with crystal-less operation, a DMIC subsystem for connection of two PDM microphone and two I2S interfaces.

Various low power features make the LPC5411x an ideal candidate for always-on sensor applications, providing on one hand very low standby current and on the other hand high processing performance in active mode.

The DMIC interface is designed to connect to digital PDM microphones and convert the PDM audio stream in hardware to 16-bit PCM audio data. The PCM data can then be used for further processing like simple mixing or filtering, bridging to the PC with USB or even perform a sophisticated task like speech recognition.

This application note explains the basic technology for voice detection and speech pattern recognition and describes the software and hardware setup for the LPC5411x voice recognition demo.

## 2. LPC5411x DMIC System

---

### 2.1 Digital microphone interface

There are two types of digital microphone in the market with different output formats, I2S and Pulse Density Modulation (PDM). The output of a PDM microphone is a 1-bit high sample rate stream which is the direct output of the Sigma-Delta modulator in the microphone. Another kind of microphone that supports I2S output includes a decimation module and various filters, it can output low frequency PCM signals.

The PDM microphone requires only two signals (clock and data) for operation, while the I2S requires three (clock, data and word select). Also, the PDM microphone has a smaller footprint than the I2S microphone.

The DMIC interface in the LPC5411x is a hardware module for directly connecting PDM digital microphone. It can sample and convert data from up to two digital PDM microphones with a selectable clock rate.

#### 2.1.1 PDM to PCM conversion

The audio data may be further processed by downstream software algorithms. The PDM data acquired from the microphone should be converted to PCM data as most algorithms support PCM as input. For power saving reasons this has been implemented in hardware. There is no Cortex-M4 CPU intervention for this conversion.

Since the PDM signal is oversampled, when converting PDM to PCM the sample rate needs to be divided by the oversampling factor; that operation is called decimation. As illustrated in the LPC5411x User Manual, the decimation operation is implemented by CIC filters.

### 2.1.2 Half band decimator and DC block filter

By using a half band decimator, the DMIC interface can provide 2 x FS PCM out. An optional high pass filter can block DC (0Hz signals). DC, which is the average of PCM amplitude, is detrimental to and not useful for audio signal processing.

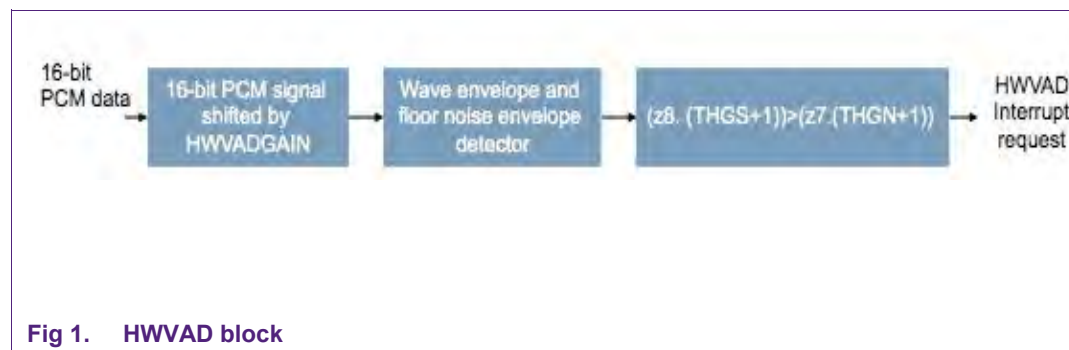
### 2.1.3 DMIC FIFO and DMA

A 16-entry hardware FIFO is implemented to accumulate audio samples. Both ARM core and DMA controller can be powered off when filling the FIFO. When FIFO is full, either an interrupt or DMA request can be triggered, per user configuration. When DMA request is enabled, data in the FIFO can be transferred to RAM by DMIC DMA without ARM core involvement. Moreover, when FIFO reaches the level specified by TRIGLVL, HWWAKE can be raised. Thus HWWAKE facilitates the gathering of DMIC audio in system low power mode.

### 2.1.4 Hardware Voice Activity Detection (HWVAD)

The HWVAD block in the LPC5411x contains a filter structure which works best for audio signals sampled with 16 kHz. Frequencies in the ultrasonic band above 16 kHz will also work as input to this block.

The HWVAD block implements an amplifier, a noise detector, a signal detector and a comparer in sequence. See [Fig 1](#). The comparer applies the threshold setting and compares the energy level of voice band to the noise energy level. HWVAD interrupt request can be raised when weighted signal energy is greater than weighted noise energy.



As mentioned in the User Manual, after reset, filters in HWVAD need time to converge, so for the first few milliseconds the data is not usable. The HWVAD interrupt should be masked on NVIV level during this time frame. If the downstream Software Voice Activity Detector (SWVAD) determines this event as invalid, then the system can go back to lowest power mode, waiting for the next HWVAD event. In case of a positive result in the SWVAD, the software enters the voice recognition phase.

## 3. Voice Recognition Technology

Voice recognition can be done in either offline or online. In online mode, LPC5411X can transfer the gathered audio data to host and eventually the speech is interpreted by server in the cloud. In offline mode, the audio data is interpreted by LPC5411X software. In this AN, the offline mode speech recognition is introduced.

Voice recognition can be classified in 2 types, speaker dependent and speaker independent. For both types, the speech recognition algorithm identifies the input voice data by analyzing the input data characteristics and processes with an existing pattern/template.

In practice, in order to save power, a voice recognition system should be in stand-by mode whenever possible as in most use cases there will be no active speech to detect for the majority of the time. This means that a convenient wake-up and timeout strategy must be implemented.

### 3.1 Speaker independent systems

In speaker independent recognition, the speech patterns/templates are preloaded in the device and no training is involved; end users are not able to modify the patterns/templates. These fixed patterns/templates are developed by the algorithm supplier using a large dataset of voice samples. Speaker independent recognition is the only option for applications where there are large number of different users, for example, ticket dispensers or elevators. Depending on the needs of the application, the training may be set up to err on the side of more false accepts (incorrect triggering based on the wrong phrase) or more false rejects (incorrect rejection of the correct phrase.) The libraries used in the demo applications referenced in this document have been tuned for typical application parameters.

Recognition performance over distance and in various signal-to-noise (SNR) conditions is affected (amongst other factors) by the sample rate of the DMIC and amount of algorithmic processing (memory and processor workload or MIPS) applied. Sample rate and processing required impact power, and thus will need to be tuned to obtain the optimal power consumption/performance for any given system. The demo code referenced in this document is optimized for the lowest power consumption and therefore, does not have the best voice recognition capability that could be achievable by a combination of the LPC5411X and Sensory's technology.

#### 3.1.1 Trigger + Commands strategy

There are two approaches for recognizing voice trigger phrases; one-step or two-steps approaches.

In a one-step approach, the system listens for all phrases all of the time. A well selected and trained trigger could work with a one-step approach, as a single command gets executed once identified. However, since speaker-independent software is less accurate, user frustration could be an issue as false detection rate vs. false acceptance rate can be very challenging to get right.

In a two-step approach, the system listens until one trigger phrase is recognized, then listens for a large set of commands. The trigger phrase can be more rigorously trained and can be more specific to the end product; for example the trigger phrase could be "Acme Thermostat", with phrases such as "make it hotter" used as commands. Phrases such as "make it hotter" could be difficult to train for a user experience in a one-step system, as similar phrases could appear in normal conversation. If a command is immediately following a trigger phrase, a command with a much lower false reject rate can be used so the end user will not get frustrated with the system not recognizing the command they are trying to give. Power savings in a two-step system are also

significant, scaling with the number of commands the system needs to be able to recognize.

### 3.2 Speaker dependent systems

Speaker-dependent software identifies a voice and phrase combination, with a focus on the unique characteristics of a single person's voice (vocal tract and individual pronunciation). The algorithm needs to be trained on specific phrases or even free selectable expressions by the individual user. This is a convenient solution for personal equipment like smartphones, fitness devices or home appliance products.

Since the end user needs to enroll the voice samples to generate new patterns/templates, end-user tools are required for the training process. Training is typically a computationally intensive task, so typically requires a higher performance host device. In that case, the host device could transmit the user-trained pattern/template to embedded device (e.g. LPC5411X) after training. Next, the pattern/template can be programmed into FLASH for offline recognition.

## 4. LPC5411x Voice recognition demo suite

LPCOpen v3 for LPC5411x is used as software framework for this demo. Source code can be found from the link below and this includes drivers for the DMIC and HWVAD:

<http://www.nxp.com/products/software-and-tools/hardware-development-tools/lpcxpresso-boards/lpcopen-software-development-platform-lpc5411x:LPCOPEN-SOFTWARE-FOR-LPC5411X?fsrch=1&sr=1&pageNum=1>

The application described in this section manages the HWVAD block and the DMIC interface in combination with third-party voice recognition (voice trigger/phrase spotting) software from Sensory (<http://www.sensory.com/>) to provide a demo voice recognition system. Source code and project files for the application (for use with Keil uVISION tools) using the "DISTYSALES" library is available for qualified customers who have NDAs with NXP and Sensory.

The performance of the demos with regards to power consumption, the microphone sensitivity as well as the tolerance for keyword recognition depends on parameters in the software framework, the hardware settings and the way the grammar files have been built.

The demos show the principle capabilities of the LPC5411x platform for this type of application.

The Sensory demo setup focuses on:

- Lowest power consumption.
- Keyword + command (two-step) approach.

### 4.1 System introduction

Logically, there are 3 components in the voice triggering system, digital microphone, DMIC interface in LPC5411x and Voice Recognition Algorithm.

As described in Fig 2, PDM microphone connects LPC5411x through DMIC interface. Pin P0\_31 is configured as PDM Channel 0 clock and pin P1\_0 is configured as PDM

Channel 0 Data-in. As of digital microphone, SEL pin is pulled to high, thus, microphone asserts data on rising clock edge and latches data on falling clock edge. The DMIC clock speed decides the oversampling ratio, which has influence on the quality of audio signal acquired and the overall system power consumption. Developers should understand the microphone characteristic and algorithm specification to decide the clock rate.

DMIC interface retrieves the PDM samples and then demodulates. Next, PCM data should be stored in a FIFO for next stage recognition algorithm. DMA interrupt and HWVAD interrupt can be raised, in sequence, then the recognition algorithm can process the audio data when there is HWVAD interrupt or DMA interrupt.

Voice triggering algorithm accepts 16 kHz PCM audio signal in, evaluates the audio data and makes a judgement if the incoming speech matches the stored pattern(s) or not.

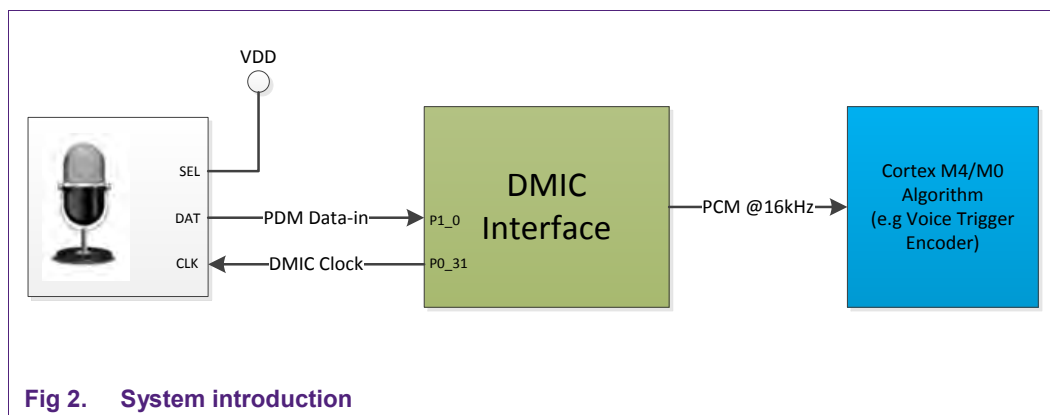


Fig 2. System introduction

## 4.2 Control flow and audio data flow

Fig 3 elaborates on the system design.

Audio data flow is marked by a green line. The step-by-step audio data routing is marked with number in sequence:

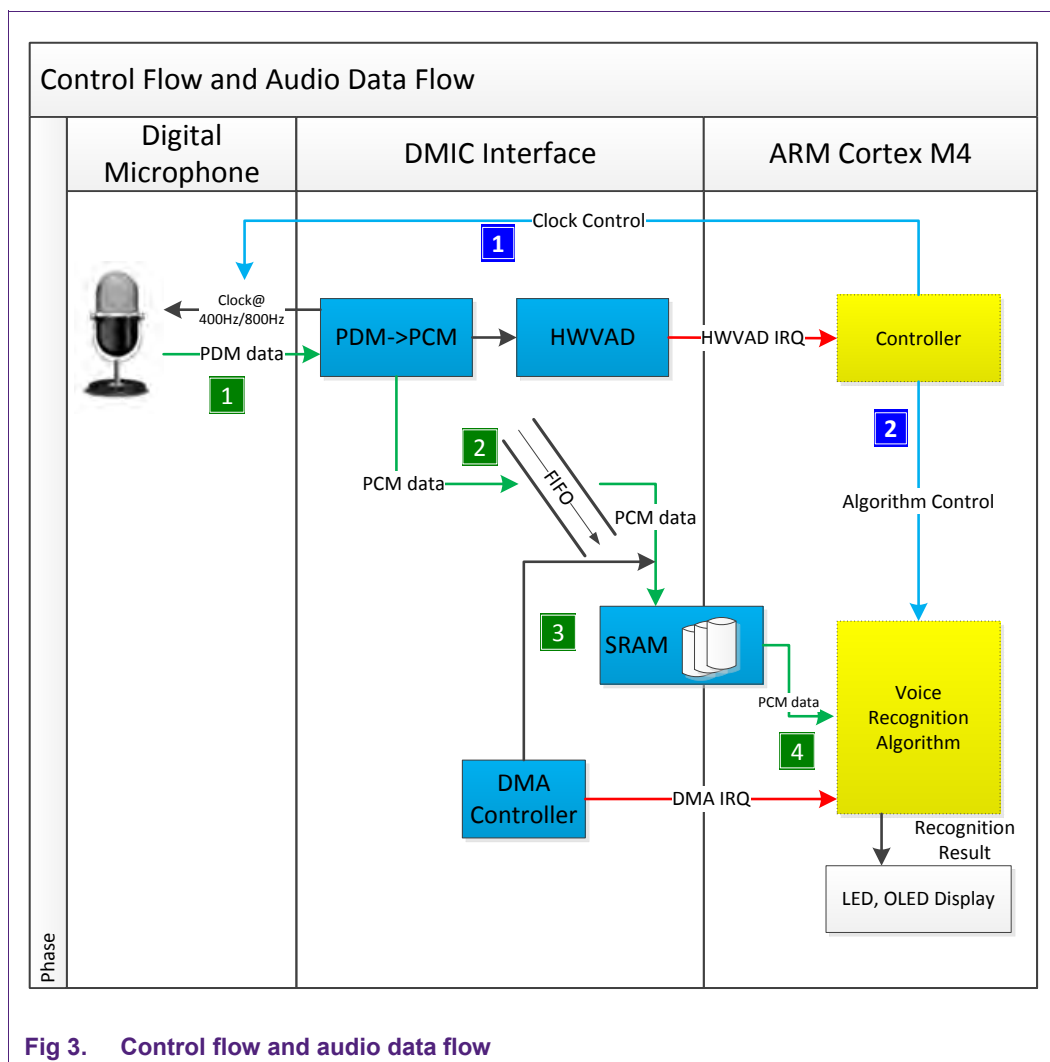
- 1) Digital microphone acquires the audio samples and encodes the data in PDM format.
- 2) DMIC interface converts the PDM data to 16 kHz PCM and stores the data in DMIC FIFO (maximum sixteen entries). The FIFO reaches the threshold in every 1 ms.
- 3) DMA Controller can move the data from FIFO to SRAM with no CPU involvement.
- 4) Voice trigger algorithm process the accumulated PCM data in every 25 ms per algorithm requirement.

Meanwhile, the system applies two interrupts, HWVAD IRQ and DMIC DMA IRQ. HWVAD IRQs are raised when voice activity is detected. As shown in Fig 3, two control signals depend on the HWVAD event:

- 1) DMIC clock control. DMIC interface supports using half rate sampling mode. When there is no voice activity, DMIC clock can be run at half rate. In half rate sampling mode, PDM data is sampled twice per clock, and the decimation rate can keep

unchanged. Using half rate sampling, power consumption is reduced without a decimator configuration change.

- 2) Algorithm control: When there is no voice activity, main clock is configured as 12 MHz and voice recognition algorithm is skipped. When there is a HWVAD event, the main clock is set to 48 MHz with FROHF enable.



### 4.3 Advanced features

Advanced features supported in the application can be enabled or disabled by switching the macro definition in C Code. The pre-built demo binaries for each command set is available at the NXP website (<http://www.nxp.com/demoboard/OM13090>), compiled for the OLED display configuration.



There are three trigger-command sets supported in the application. Developers can select the preferred set in the project configuration. The triggers and commands are introduced in [Table 1](#).

Command suites	Trigger	Commands
DISTYSALES	Hello Blue Genie	<ul style="list-style-type: none"> <li>- NXP LOGO</li> <li>- Sensory LOGO</li> <li>- Make it so</li> <li>- Blink LED</li> <li>- Flash display</li> </ul>
Home Automation		<ul style="list-style-type: none"> <li>- Turn on the Light</li> <li>- Turn the Light off</li> <li>- Turn on the Air conditioner</li> <li>- Turn the Air Conditioner off</li> <li>- Open Garage Door</li> <li>- Close Garage</li> <li>- What's the temperature</li> <li>- Kitchen Turn on Light</li> <li>- Kitchen turn the Light off</li> </ul>
Thermostat		<ul style="list-style-type: none"> <li>- Make it Colder</li> <li>- Make it Cooler</li> <li>- Make it Hotter</li> <li>- Make it Warmer</li> <li>- Outside Temperature</li> <li>- Room Temperature</li> <li>- Today's Weather</li> </ul>

**Table 1. Trigger command sets**

### USE\_LEDSIGNALS/OLED

There are two human machine interface (HMI) options for recognition event notification. One is OLED display, another is three-color LEDs. When the specific speech pattern is matched, OLED can display different preload pictures with OLED enabled. Similarly, LED

can blink differently according to the pattern matched when USE\_LEDSIGNALS is enabled. More details can be found in section 5.3.

#### **HWVAD\_OFF**

By default, the hardware VAD module of DMIC is enabled. HWVAD can raise an interrupt when there is a voice band activity.

#### **TRIG\_AND\_CMD/MULTI\_TRIG**

TRIG\_AND\_CMD should be enabled in the application. It enable the 2-steps trigger + command recognition.

#### **STEREO\_PCM/MONO\_PCM**

As there is only one MIC on the MAO board, the STEREO\_PCM should be disabled and MONO\_PCM should be enabled. In general, audio channel number determines the audio buffer size. From software aspect, configuration of DMIC configuration and DMA configuration are differ with different channel numbers.

#### **DMIC\_CLOCK\_HALF**

When DMIC\_CLOCK\_HALF is enabled, the DMIC clock is running at half rate when there is no voice activity. Lower DMIC clock results to low power consumption of DMIC. When there is voice activity DMIC clock is raised to full speed, so that higher quality of PCM signal can be acquired for downstream algorithm.

#### **MULTI\_BUFFER**

When MULTI\_BUFFER is enabled, 8 buffers are declared between the DMIC output and downstream algorithm, acting as FIFO. Since the recognition algorithm may require more time in a specific step, the FIFO ensures the data integrity and no audio data loss.

#### **USE\_WDT\_OSC\_FOR\_DMIC**

Two clock sources are selectable for the DMIC clock: watchdog oscillator or internal free running oscillator (FRO). The Watch Dog oscillator has an output range from 200 kHz to 1.5MHz. The working current of this oscillator is 2uA while 100uA is required for the FRO; thus setting USE\_WDT\_OSC\_FOR\_DMIC can save 98uA system current. However, the watchdog oscillator is not accurate when compared to the FRO. The accuracy of WDT is limited to +/-40% over temperature, voltage, and silicon processing variations, so can impair system performance when a wide operating conditions are expected. The DMIC sampling time can also have jitter and delay when using WDT as DMIC clock source.

#### **STREAM\_PCM**

There is a means to dump the acquired PCM data in runtime through UART. The UART is running at baud rate 921600 in blocking mode. As the dumped PCM is the direct input of voice recognition algorithm, thus, PCM dumping is an important measurement to enhance the recognition rate. At the host computer side PCM signal can be further analyzed in time and frequency domain. Note that it may be necessary to use an external USB to serial cable connected to the P3 connector of the LPCXpresso54114 board in order to sustain this data rate (as opposed to using the built-in VCOM feature of the board).

### GPIO\_MARKER

Three PINs (P0\_18, P0\_19, and P0\_20) are configured as output GPIOs for debugging and triggering external actions when voice triggers are detected. The feature should be disabled when performing power measurement.

## 5. LPC5411x demo description

### 5.1 Voice triggering demo suite

OM13090 kit for LPC5411x

- LPCXpresso54114 board with on-board debugger bridge.
- MAO Shield board with DMIC, OLED display, and audio codec.

For more details visit the following link:

<http://www.nxp.com/products/software-and-tools/hardware-development-tools/lpcxpresso-boards/lpc54114-audio-and-voice-recognition-kit:OM13090>

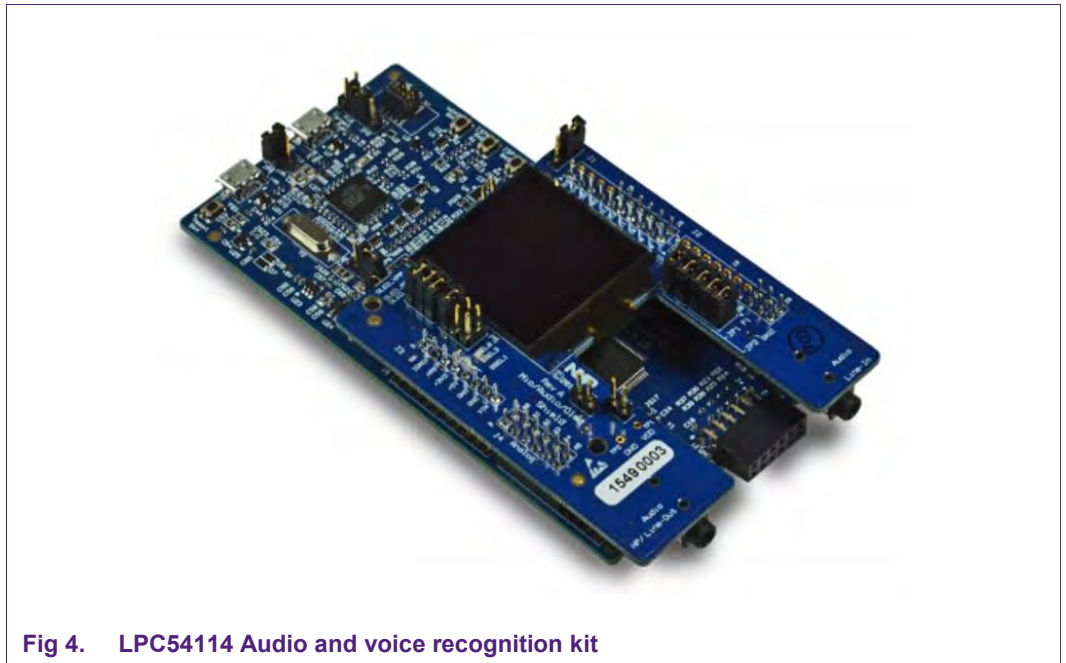
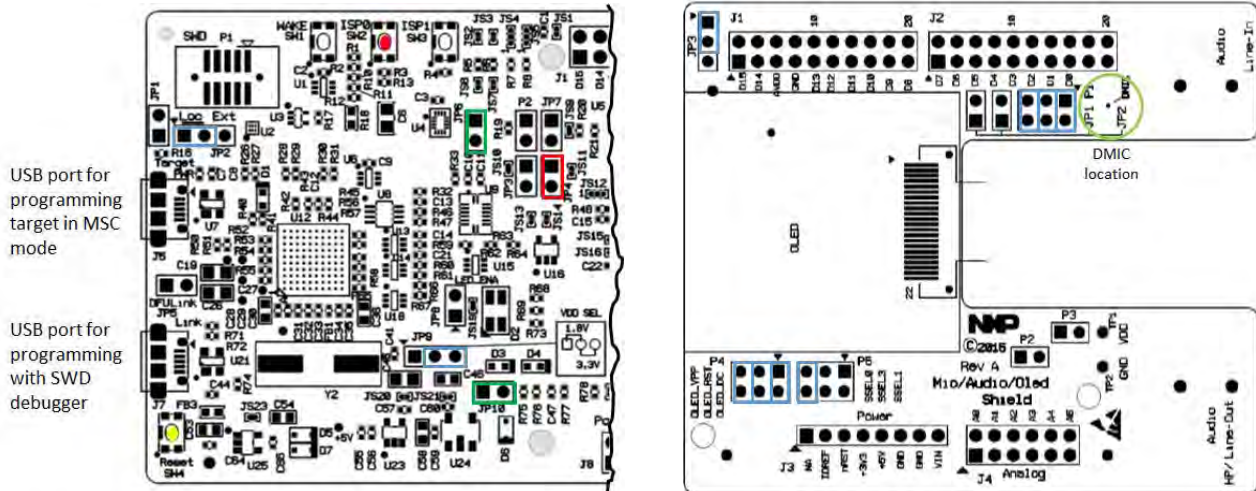


Fig 4. LPC54114 Audio and voice recognition kit

## 5.2 Hardware setup for binary programming

The LPC54114 supports flash programming using USB mass storage mode. This feature provides a simple and quick way to load different binaries onto the Kit. The figure below shows how to configure the board for MSC programming.



- Configure the board as shown
- Connect USB port J5 to PC
- Press and hold ISP0 and press Reset button
- Setting the red jumper is only required in case JS11 has been taken out for enabling current measurements over the pins of JP4

- Configure the OLED shield as shown above for correct OLED and DMIC operation

Fig 5. Configuration of OM13090 for MSC programming

1. The PC should show a mass storage device called CRP\_DISABLED. The drive is shown with a size of 266.244 bytes. However, the physical size of the flash is 256 KB. This is due to a virtual block of 4 KB used by the mass storage firmware.
2. Open the MSC device with a file manager and delete the file *firmware.bin*. After this action the file manager will show no more files, but the flash still has a fully occupied memory. This differs from a standard mass storage device, such as a USB stick.
3. Drag and drop the new binary file to the drive. Note that this new file **must** have the name *firmware.bin*, otherwise the file seems to be transferred and programmed but in fact it is not.
4. Reset the board.

## 5.3 Sensory demo

The demo with the Sensory voice recognition solution is working with a two-step approach, taking DISTYSALES build for example, just by definition

**Keyword:** *Hello Blue Genie*

<b>Commands:</b>	<i>Make it so</i>	(1)
	<i>NXP logo</i>	(2)
	<i>Flash display</i>	(3)
	<i>Sensory logo</i>	(4)
	<i>Blink L E D</i>	(5)

After saying the keyword there is a specific time frame for command recognition. This time can be defined by the application programmer. In case a command is said and recognized by the MCU, the software will show feedback on the OLED display and then go back to deep sleep mode. When there is no command or the command is not understood, the software goes back to deep sleep mode after a timeout.



## 6. Power consumption considerations

A product with voice recognition normally needs to be always-on and the system must be able to recognize a keyword/command at any time. For tight energy budget /battery products, the LPC5411x offers various methods to save power, such as flexible clocking options or power save modes with fast wake-up.

### 6.1 Hardware setup for power measurement

The LPC5411x LPCXpresso board has an on-board measurement amplifier which allows to measure the power in combination with the LPCXpresso IDE. See documentation under the LPCXpresso IDE:

<http://www.nxp.com/lpcxpressoide>

For lower currents below ~300-μA this amplifier will introduce a noticeable error due to its input voltage offset. See **Error! Reference source not found..**

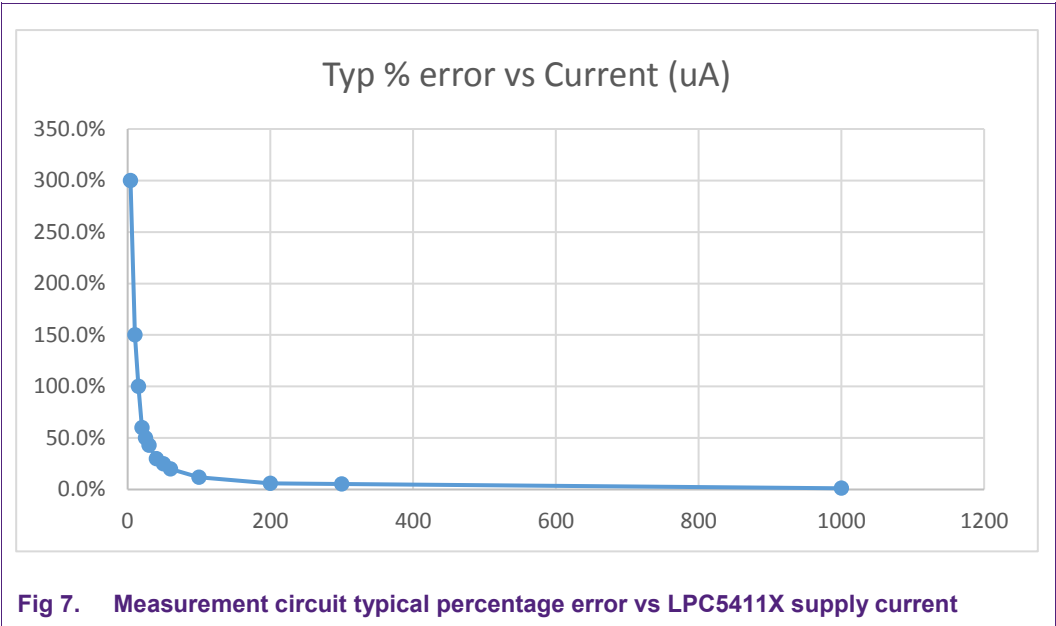


Fig 7. Measurement circuit typical percentage error vs LPC5411X supply current

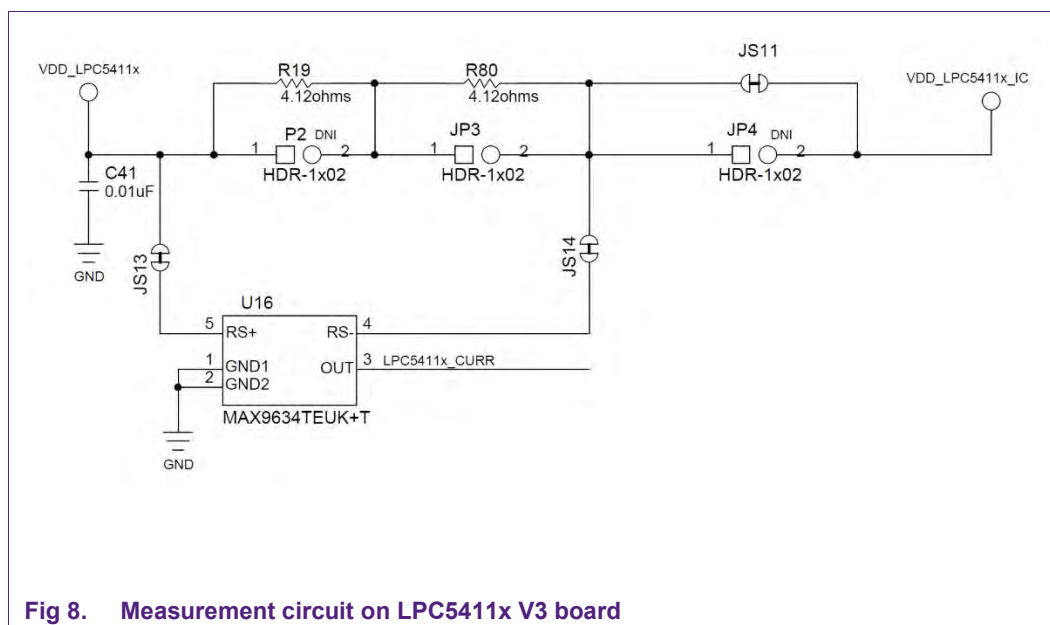
However, it is also possible to use a standard ammeter supporting μA for measuring the power consumption of the MCU. For more details, see the application note, [AN11799: LPC5411x Low Power Modes and Wake-up Times](#).

Table 2. LPCXpresso board modifications for power measurements using ampere-meter	
Modification	Description
Open JS11	JS11 is assembled with a 0 Ohm resistor → remove it
Solder pins in JP4	Connect the ampere-meter across pins of JP4
Solder pins in P2 and JP3	With jumpers on P2 and JP3 the 4.12 Ohm shunts can be bypassed

Alternatively (without doing any hardware modifications) the MCU current can also be measured with a volt-meter over the 8.24 Ohm shunt (R19 + R80), calculating the current as follows:

$$I = \frac{U}{8.240hm}$$

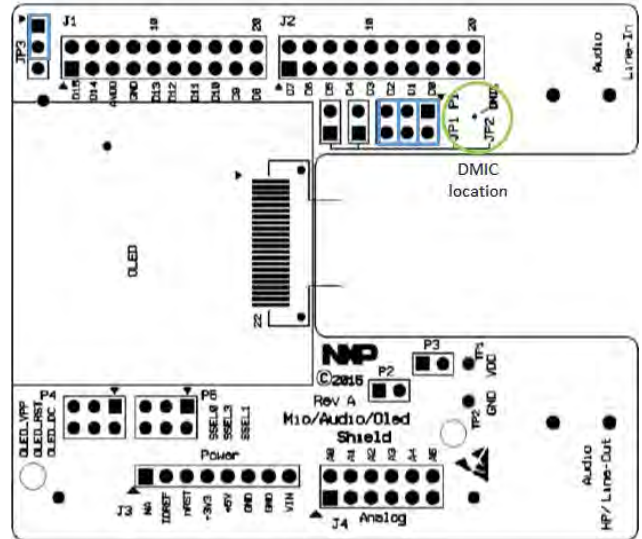
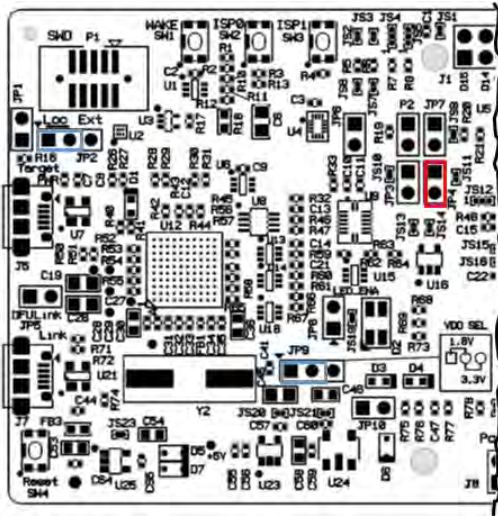
For this the volt-meter is connected to pins of P2 and JP3.



The Sensory demo source code comes in two different versions (select the files from their respective folders). One is supporting the OLED shield (folder Sensory OLED) and consumes some MCU power for the OLED I/O. The other version in folder Sensory LED gives feedback using the on-board RGB-LED and uses just the DMIC on the shield board. For lowest power operation, the I/O voltage must be set to 1.8 V and the connection to the OLED display must be cut. See [Fig 9](#).



USB port  
for power



- Configure the board as shown to 1.8V I/O
- Connect USB power to port J5
- Current measurement is done on the red jumper location JP4

- Configure the OLED shield as shown above for DMIC only operation

Fig 9. Configuration of OM13090 for low power measurement

Fig 10 shows how the system behaves in this LED configuration. After reset the software goes through the MCU initialization and enters listening mode. If no keyword ("Hello Blue Genie") is recognized the green LED flashes one time to indicate the transition into lowest possible power mode (deep-sleep mode with audio data batching, waiting for HWVAD event). After a HWVAD event the software goes to listening mode, waiting for the keyword. With a successful keyword recognition the software goes to command mode, indicated by the blue LED flashing two times. Otherwise a timeout occurs, indicated by the green LED. The timeout period needs to be specified by the application programmer according to the product requirements. After a command is recognized, the red LED flashes n times, indicating the command number (see section 5.3 for commands), followed by a transition back to lowest power mode.



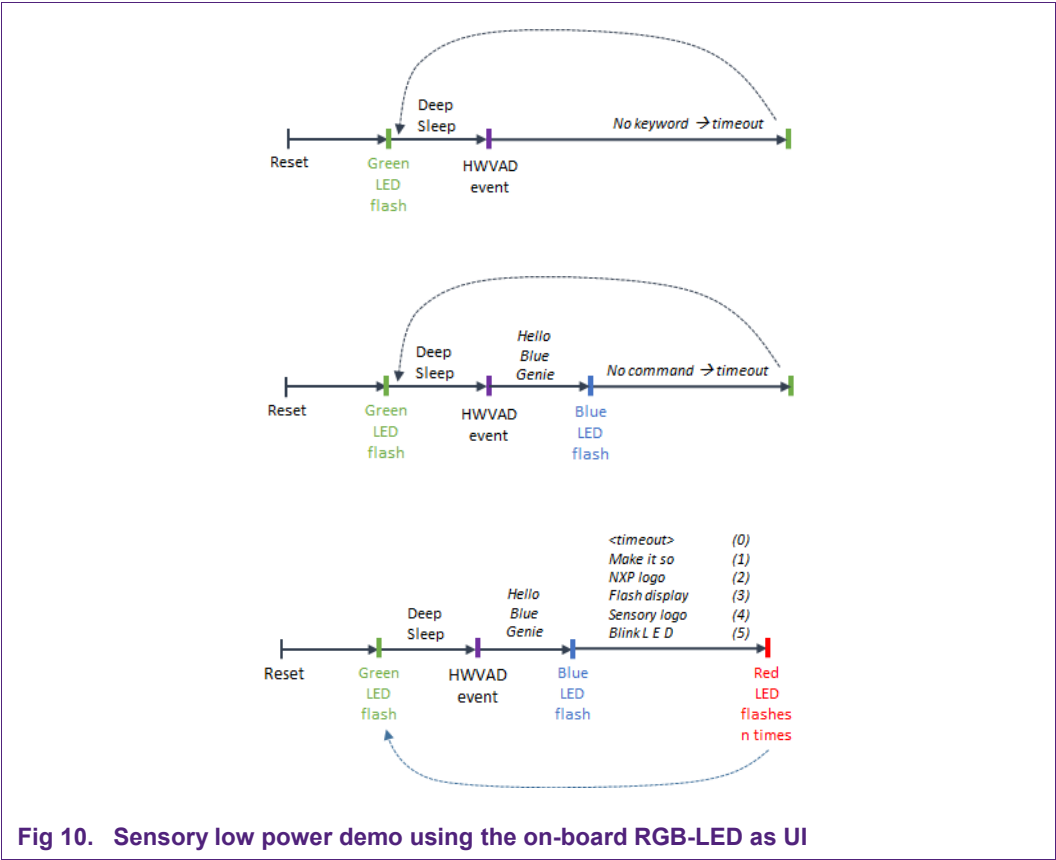


Fig 10. Sensory low power demo using the on-board RGB-LED as UI

Fig 11 shows the average power consumption during the different stages of the software.

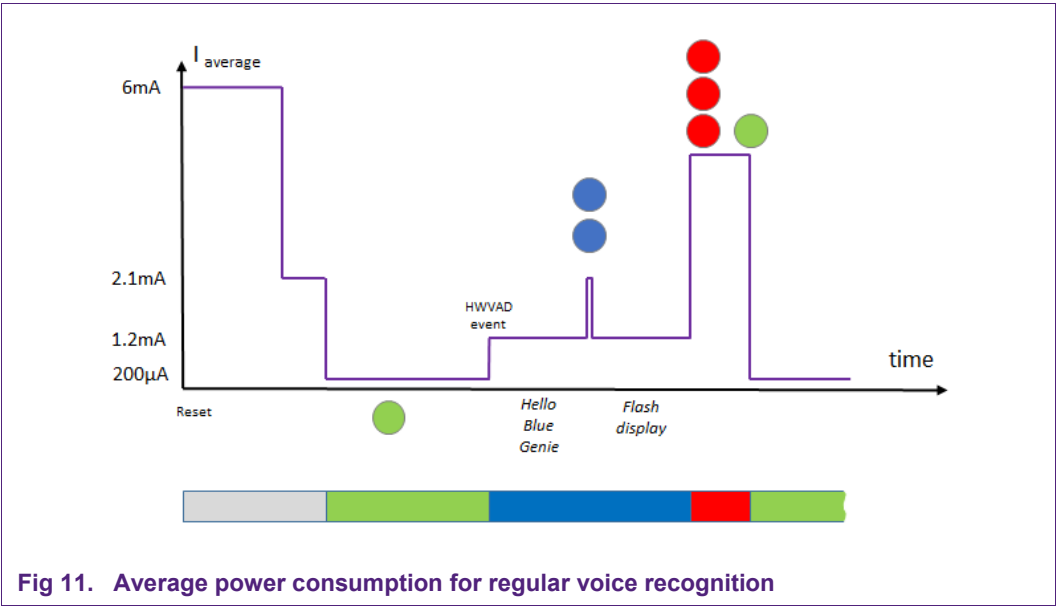


Fig 11. Average power consumption for regular voice recognition

**Reset and initialization phase:**

- Generic LPCOpen startup and initialization of the CPU clock to 48MHz FRO clock, followed by application code applying power saving settings and a reduction to 12MHz CPU clock.

**Listening phase:**

- Audio data sampling at 400 kHz DMIC clock, data stored into SRAM using DMA.
- No Cortex-M4 activity, MCU is in deep-sleep mode.

**Recognition phase:**

- HWVAD interrupt woke up the MCU.
- DMIC clock frequency increased to 800 kHz.
- Cortex-M4 runs at 48 MHz using FRO as clock source and analyzes the incoming audio data from time to time (chunks of 240 bytes). In between it goes back to deep sleep mode.

**Signaling phase:**

- Cortex-M4 runs at 48MHz using FRO as clock source, busy with some UI signaling, without going to sleep or deep sleep mode.

**Table 3. Power consumption results for Sensory demo**

Demo type	Reset / Init	Listening	Recognition	Signaling
Sensory OLED	6 mA	460 $\mu$ A	1.4 mA	4.6 mA
Sensory LED	6 mA	92 $\mu$ A	1 mA	4.2 mA

## 7. Legal information

### 7.1 Definitions

**Draft** — The document is a draft version only. The content is still under internal review and subject to formal approval, which may result in modifications or additions. NXP Semiconductors does not give any representations or warranties as to the accuracy or completeness of information included herein and shall have no liability for the consequences of use of such information.

### 7.2 Disclaimers

**Limited warranty and liability** — Information in this document is believed to be accurate and reliable. However, NXP Semiconductors does not give any representations or warranties, expressed or implied, as to the accuracy or completeness of such information and shall have no liability for the consequences of use of such information.

In no event shall NXP Semiconductors be liable for any indirect, incidental, punitive, special or consequential damages (including - without limitation - lost profits, lost savings, business interruption, costs related to the removal or replacement of any products or rework charges) whether or not such damages are based on tort (including negligence), warranty, breach of contract or any other legal theory.

Notwithstanding any damages that customer might incur for any reason whatsoever, NXP Semiconductors' aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms and conditions of commercial sale of NXP Semiconductors.

**Right to make changes** — NXP Semiconductors reserves the right to make changes to information published in this document, including without limitation specifications and product descriptions, at any time and without notice. This document supersedes and replaces all information supplied prior to the publication hereof.

**Suitability for use** — NXP Semiconductors products are not designed, authorized or warranted to be suitable for use in life support, life-critical or safety-critical systems or equipment, nor in applications where failure or malfunction of an NXP Semiconductors product can reasonably be expected to result in personal injury, death or severe property or environmental damage. NXP Semiconductors accepts no liability for inclusion and/or use of NXP Semiconductors products in such equipment or applications and therefore such inclusion and/or use is at the customer's own risk.

**Applications** — Applications that are described herein for any of these products are for illustrative purposes only. NXP Semiconductors makes no representation or warranty that such applications will be suitable for the specified use without further testing or modification.

Customers are responsible for the design and operation of their applications and products using NXP Semiconductors products, and NXP Semiconductors accepts no liability for any assistance with applications or customer product design. It is customer's sole responsibility to determine whether the NXP Semiconductors product is suitable and fit for the customer's applications and products planned, as well as for the planned application and use of customer's third party customer(s). Customers should provide appropriate design and operating safeguards to minimize the risks associated with their applications and products.

NXP Semiconductors does not accept any liability related to any default, damage, costs or problem which is based on any weakness or default in the

customer's applications or products, or the application or use by customer's third party customer(s). Customer is responsible for doing all necessary testing for the customer's applications and products using NXP Semiconductors products in order to avoid a default of the applications and the products or of the application or use by customer's third party customer(s). NXP does not accept any liability in this respect.

**Export control** — This document as well as the item(s) described herein may be subject to export control regulations. Export might require a prior authorization from national authorities.

**Evaluation products** — This product is provided on an "as is" and "with all faults" basis for evaluation purposes only. NXP Semiconductors, its affiliates and their suppliers expressly disclaim all warranties, whether express, implied or statutory, including but not limited to the implied warranties of non-infringement, merchantability and fitness for a particular purpose. The entire risk as to the quality, or arising out of the use or performance, of this product remains with customer.

In no event shall NXP Semiconductors, its affiliates or their suppliers be liable to customer for any special, indirect, consequential, punitive or incidental damages (including without limitation damages for loss of business, business interruption, loss of use, loss of data or information, and the like) arising out of the use of or inability to use the product, whether or not based on tort (including negligence), strict liability, breach of contract, breach of warranty or any other theory, even if advised of the possibility of such damages.

Notwithstanding any damages that customer might incur for any reason whatsoever (including without limitation, all damages referenced above and all direct or general damages), the entire liability of NXP Semiconductors, its affiliates and their suppliers and customer's exclusive remedy for all of the foregoing shall be limited to actual damages incurred by customer based on reasonable reliance up to the greater of the amount actually paid by customer for the product or five dollars (US\$5.00). The foregoing limitations, exclusions and disclaimers shall apply to the maximum extent permitted by applicable law, even if any remedy fails of its essential purpose.

### 7.3 Licenses

#### Purchase of NXP <xxx> components

<License statement text>

### 7.4 Patents

Notice is herewith given that the subject device uses one or more of the following patents and that each of these patents may have corresponding patents in other jurisdictions.

<Patent ID> — owned by <Company name>

### 7.5 Trademarks

Notice: All referenced brands, product names, service names and trademarks are property of their respective owners.

<Name> — is a trademark of NXP B.V.

## 8. List of figures

---

Fig 1.	HWVAD block .....	4
Fig 2.	System introduction .....	7
Fig 3.	Control flow and audio data flow .....	8
Fig 4.	LPC54114 Audio and Voice Recognition Kit...	11
Fig 5.	Configuration of OM13090 for MSC programming .....	12
Fig 6.	Sensory demo example .....	13
Fig 7.	Measurement circuit typical percentage error vs LPC5411X supply current .....	14
Fig 8.	Measurement circuit on LPC5411x V3 board..	15
Fig 9.	Configuration of OM13090 for low power measurement .....	16
Fig 10.	Sensory low power demo using the on-board RGB-LED as UI.....	17
Fig 11.	Average power consumption for regular voice recognition .....	17

9. List of tables

Table 1. Trigger Command Sets.....9

Table 2. LPCXpresso board modifications for power  
measurements using ampere-meter ..... 14

Table 3. Power consumption results for Sensory demo 18

## 10. Contents

---

<b>1.</b>	<b>Introduction .....</b>	<b>3</b>
<b>2.</b>	<b>LPC5411x DMIC System .....</b>	<b>3</b>
2.1	Digital microphone interface.....	3
2.1.1	PDM to PCM conversion.....	3
2.1.2	Half band decimator and DC block filter .....	4
2.1.3	DMIC FIFO and DMA.....	4
2.1.4	Hardware Voice Activity Detection (HWVAD).....	4
<b>3.</b>	<b>Voice Recognition Technology .....</b>	<b>4</b>
3.1	Speaker independent systems .....	5
3.1.1	Trigger + Commands strategy.....	5
3.2	Speaker dependent systems.....	6
<b>4.</b>	<b>LPC5411x Voice recognition demo suite .....</b>	<b>6</b>
4.1	System introduction.....	6
4.2	Control flow and audio data flow .....	7
4.3	Advanced features .....	8
<b>5.</b>	<b>LPC5411x demo description .....</b>	<b>11</b>
5.1	Voice triggering demo suite.....	11
5.2	Hardware setup for binary programing.....	12
5.3	Sensory demo .....	12
<b>6.</b>	<b>Power consumption considerations.....</b>	<b>13</b>
6.1	Hardware setup for power measurement .....	13
<b>7.</b>	<b>Legal information .....</b>	<b>19</b>
7.1	Definitions .....	19
7.2	Disclaimers.....	19
7.3	Licenses.....	19
7.4	Patents.....	19
7.5	Trademarks.....	19
<b>8.</b>	<b>List of figures.....</b>	<b>20</b>
<b>9.</b>	<b>List of tables .....</b>	<b>21</b>
<b>10.</b>	<b>Contents.....</b>	<b>22</b>

---

Please be aware that important notices concerning this document and the product(s) described herein, have been included in the section 'Legal information'.

---