

# 使用 TLV320ADCx120 和 PCMx120-Q1 中的语音活动检测器 (VAD)



Muthuraj Pillai, Diljith Thodi

Low Power Audio

## 摘要

TLV320ADCx120/PCMx120-Q1 系列器件 ( TLV320ADC3120/PCM3120-Q1、TLV320ADC5120/PCM5120-Q1 和 TLV320ADC6120/PCM6120-Q1 ) 是用于音频应用的双通道、高性能模数转换器 (ADC)。该系列器件具有广泛的功能，其中包括：

- 可编程通道增益
- 数字音量控制
- 可编程麦克风偏置电压
- 锁相环 (PLL)
- 可编程高通滤波器 (HPF)
- 自动增益控制 (AGC)
- TLV320ADC5120/PCM5120-Q1 和 TLV320ADC6120/PCM6120-Q1 中的动态范围增强器 (DRE) 支持
- 支持高达 768kHz 采样率的线性相位或低延迟滤波器模式
- 语音活动检测器 (VAD)

本应用手册介绍了如何在 TLV320ADCx120/PCMx120-Q1 器件中配置语音活动检测器 (VAD) 功能。

## 内容

1 引言.....	2
2 语音活动检测器.....	3
2.1 VAD 配置.....	4
2.2 VAD 参数.....	6
3 VAD 结果.....	8
4 示例.....	10
5 相关文档.....	11
6 修订历史记录.....	11

## 插图清单

图 1-1. VAD 示例.....	2
图 2-1. VAD 方框图.....	3
图 3-1. 针对汽车噪声的非语音命中与语音命中.....	8
图 3-2. 针对餐厅噪声的非语音命中与语音命中.....	8
图 3-3. 针对地铁噪声的非语音命中与语音命中.....	9
图 3-4. 针对 12dB SNR 在 -7dB 阈值下的非语音命中与语音命中.....	9

## 表格清单

表 2-1. VAD 配置列表.....	4
表 2-2. 使用 VAD_CFG1 寄存器选择 VAD 模式.....	4
表 2-3. 使用 VAD_CFG1 寄存器选择 VAD 通道.....	4
表 2-4. 使用 VAD_CFG1 寄存器选择 VAD 时钟.....	5
表 2-5. 使用 VAD_CFG1 寄存器选择 VAD 时钟频率.....	5
表 2-6. 使用 VAD_CFG2 寄存器选择在记录期间 VAD 开启.....	5

表 2-7. 使用 VAD_CFG2 寄存器选择 SDOUT 用作中断.....	5
表 2-8. VAD 参数列表.....	6
表 2-9. 用于初始学习周期的可编程系数寄存器.....	6
表 2-10. 用于保持计数器的可编程系数寄存器.....	6
表 2-11. 用于唤醒等待的可编程系数寄存器.....	7
表 2-12. 用于阈值的可编程系数寄存器.....	7

## 商标

所有商标均为其各自所有者的财产。

## 1 引言

语音活动检测器 (VAD) 是一种语音触发的系统唤醒机制。VAD 使系统的其余部分能够在没有语音活动的情况下处于睡眠模式，从而消耗非常低的功率。基于 VAD 的系统在检测到语音活动时会产生中断。图 1-1 显示了 VAD 如何响应语音活动。



图 1-1. VAD 示例

TLV320ADCx120/PCMx120-Q1 器件系列的所有模数转换器 (ADC) 通道 (即模拟和数字麦克风通道) 都支持 VAD。对于低功耗应用，优先使用数字麦克风通道。本应用手册介绍了 VAD 的运行情况、可调参数以及支持 VAD 所需的器件配置。

## 2 语音活动检测器

VAD 块监视来自麦克风通道的信号，以寻找类似语音的模式，并在检测到匹配模式时触发中断。VAD 监视语音活动的开始和语音活动的结束。这两个事件都可以映射到中断。

ADCx120 器件还能够根据 VAD 中断自动开启和关闭电源。例如，可以将 ADCx120 系统设置为在数字麦克风通道上监视 VAD 活动，然后根据 VAD 触发开启模拟麦克风通道电源。

VAD 有两种工作模式：

1. 自动模式：VAD 自动触发，以开启和关闭系统电源。
2. 用户模式：必须由主机监视 VAD 中断，并且必须通过 I2C 命令开启或关闭系统电源。

请注意，在自动模式和用户模式下，该器件还会在 GPIO 或 GPO 引脚上生成一个中断，该中断可以发送到外部 DSP 或 SOC。

VAD 的显著特性如下：

- 无需外部时钟 - 系统重复使用经过调谐的内部振荡器，来生成具有所需频率的 ADC 和 DSP 时钟。
- 在 VAD 模式和记录模式之间自动切换 - 系统在出现语音活动时从 VAD 模式切换到记录模式，并在自动 VAD 模式下没有语音活动时切换回至 VAD 模式。
- 可通过 RAM 重新配置决策树 - 利用该特性，可以使用已针对其他声学事件检测进行训练的不同决策树来更新默认语音活动决策树。

VAD 算法使用基于决策树分级的算法进行语音活动检测。可以通过系数写入更新决策树参数，因此可以针对需要使用决策树进行检测的其他应用重新配置 VAD 块。16 频段非均匀间隔 IIR 滤波器组用于特征提取。可以离线计算特征选择参数和决策树，并通过系数存储器写入对其进行更新。图 2-1 显示了 VAD 的信号处理链。

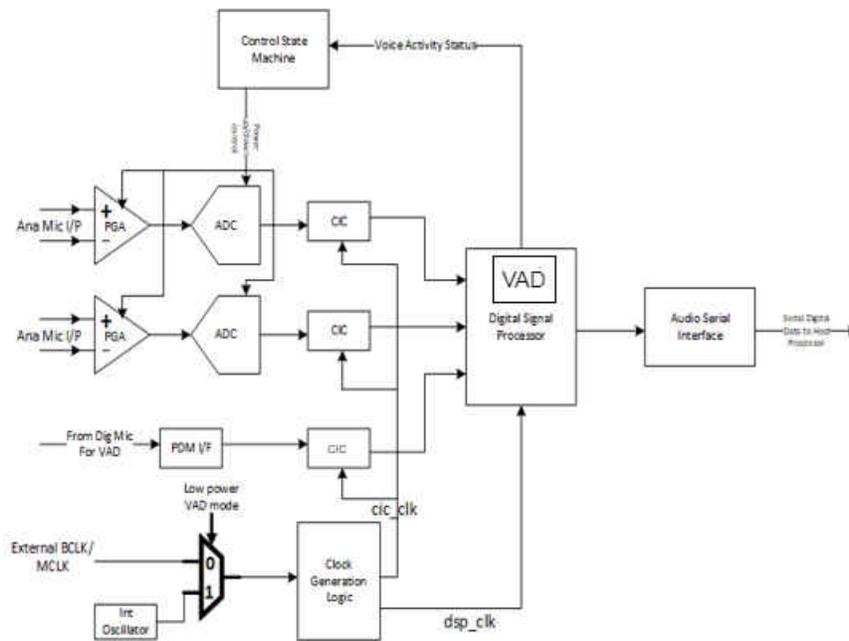


图 2-1. VAD 方框图

## 2.1 VAD 配置

表 2-1 显示了可以运行 VAD 的不同模式。

表 2-1. VAD 配置列表

VAD 配置	功能、说明
用户、自动	自动：ADC 上电和断电分别在检测到语音和无语音时自动发生。 用户：在检测到语音时，会生成中断，用户可以启动 ADC 上电和断电。
VAD 通道	决定要在哪个通道上监视 VAD 活动。
VAD 时钟	VAD 需要根据内部时钟或外部时钟运行。
在记录期间 VAD 开启	该功能决定在记录过程中语音检测是否需要处于活动状态。
SDOUT 中断	在未记录通道数据时启用 SDOUT 引脚，以支持中断输出。

**用户、自动**：用户可以将 VAD 编程为处于自动模式或用户模式。用户可以对 VAD 进行编程，使其处于 4 种可能的模式之一。

**0d = 用户启动的 ADC 上电和 ADC 断电**：用户可以根据 VAD 算法生成的中断来启动 ADC 上电和 ADC 断电。

**1d = 基于 VAD 中断的 ADC 上电和 ADC 断电**：这是自动模式，其中 ADC 根据 VAD 算法生成的中断自动开启或关闭。

如表 2-2 所示，可以使用 VAD\_CFG1 寄存器（页面 = 0x01，地址 = 0x1E）的 VAD\_mode[1:0] 位来完成 VAD 模式选择。

表 2-2. 使用 VAD\_CFG1 寄存器选择 VAD 模式

位	字段	类型	复位	说明
7-6	VAD_MODE[1:0]	读/写	00b	自动 ADC 上电和断电配置选择。 0d = 用户启动的 ADC 上电和 ADC 断电 1d = 基于 VAD 中断的 ADC 上电和 ADC 断电 2d = 基于 VAD 中断的 ADC 上电，但用户启动的 ADC 断电 3d = 用户启动的 ADC 上电，但基于 VAD 中断的 ADC 断电

**VAD 通道**：该参数决定要在哪个通道上监视 VAD 活动。一次只能在一个通道上监视 VAD 活动。

如表 2-3 所示，可以使用 VAD\_CFG1 寄存器（页面 = 0x01，地址 = 0x1E）的 VAD\_CH\_SEL[1:0] 位来完成 VAD 通道选择。

表 2-3. 使用 VAD\_CFG1 寄存器选择 VAD 通道

位	字段	类型	复位	说明
5-4	VAD_CH_SEL[1:0]	R/W	10b	VAD 通道选择。 0d = 在通道 1 上监视 VAD 活动 1d = 在通道 2 上监视 VAD 活动 2d = 在通道 3 上监视 VAD 活动 3d = 在通道 4 上监视 VAD 活动

**VAD 时钟**：VAD 可以根据内部振荡器时钟或用户提供的外部时钟运行。可以在 BCLK 引脚或 MCLK 引脚上提供该外部时钟。

如表 2-4 所示，可以使用 VAD\_CFG1 寄存器（页面 = 0x01，地址 = 0x1E）的 VAD\_CLK\_CFG[1:0] 位来完成 VAD 通道选择。如果用户选择 01b 或 10b，则可以使用 VAD\_CFG1 寄存器（页面 = 0x01，地址 = 0x1E）的 VAD\_EXT\_CLK\_CFG[1:0] 位来选择外部时钟的频率，如表 2-5 所示。

表 2-4. 使用 VAD\_CFG1 寄存器选择 VAD 时钟

位	字段	类型	复位	说明
3-2	VAD_CLK_CFG[1:0]	读/写	00b	VAD 的时钟选择 0d = 使用内部振荡器时钟的 VAD 处理 1d = 使用 BCLK 输入上的外部时钟的 VAD 处理 2d = 使用 MCLK 输入上的外部时钟的 VAD 处理 3d = 基于页面 0 中的 MST_CFG、CLK_SRC 和 CLKGEN_CFG 寄存器的自定义时钟配置

表 2-5. 使用 VAD\_CFG1 寄存器选择 VAD 时钟频率

位	字段	类型	复位	说明
1-0	VAD_EXT_CLK_CFG[1:0]	读/写	00b	使用外部时钟的 VAD 时钟配置。 0d = 外部时钟为 3.072MHz 1d = 外部时钟为 6.144MHz 2d = 外部时钟为 12.288MHz 3d = 外部时钟为 18.432MHz

**在记录期间 VAD 开启**：该参数决定在 ADC 记录过程中是否需要检测语音活动。如果启用该位，则 VAD 算法会在 ADC 记录过程中继续运行，以检测任何语音活动。

如表 2-6 所示，可以使用 VAD\_CFG2 寄存器（页面 = 0x01，地址 = 0x1F）的 VAD\_PD\_DET\_EN 位来完成在记录期间 VAD 开启选择。

表 2-6. 使用 VAD\_CFG2 寄存器选择在记录期间 VAD 开启

位	字段	类型	复位	说明
3	VAD_PD_DET_EN	R/W	1b	在 VAD 活动期间启用 ASI 输出数据。 0d = 在 ADC 记录期间不启用 VAD 处理 1d = 在 ADC 记录期间启用 VAD 处理，并按照配置生成 VAD 中断

**SDOUT 用作中断**：当未在进行 ADC 记录时，SDOUT 引脚可用于 VAD 中断。通过设置该位，使 SDOUT 可用作 VAD 中断引脚。

如表 2-7 所示，可以使用 VAD\_CFG2 寄存器（页面 = 0x01，地址 = 0x1F）的 SDOUT\_INT\_CFG 位来完成 SDOUT 用作中断选择。

表 2-7. 使用 VAD\_CFG2 寄存器选择 SDOUT 用作中断

位	字段	类型	复位	说明
6	SDOUT_INT_CFG	R/W	0b	SDOUT 中断配置。 0d = 未使 SDOUT 引脚支持中断功能 1d = 使 SDOUT 引脚在未记录通道数据期间支持中断输出

## 2.2 VAD 参数

表 2-8 显示了 VAD 算法的参数。这些参数驻留在器件的 32 位宽系数存储器 ( Book 0、Page 1、Page 7、Page 8 和 Page 9 ) 中。

表 2-8. VAD 参数列表

VAD 参数	功能、说明
初始学习周期 (ILP)	这是从 VAD 开启瞬间开始 VAD 算法调整自身以适应背景噪声环境所花费的时间。
保持计数器 (HOC)	在检测到语音活动时，VAD 算法会生成一个中断，如果该中断被编程为高电平有效，则在检测到语音时中断变为高电平 ( 逻辑 1 )，当没有语音时变为低电平 ( 逻辑 0 )。在变为低电平之前，语音活动停止后中断保持高电平的时长由 HOC 计数决定。
唤醒等待 (WW)	如果 VAD 被编程为处于自动模式，则在检测到语音时，它将自动开启 ADC 并开始记录，同时还会检查是否存在语音活动。唤醒等待是 VAD 在进入记录模式后暂停的时长，经过该时长后会恢复。
阈值 (TH)	阈值控制决策树节点的决策边界。较高的值将增加决策树所有节点的节点阈值，从而降低误报的可能性。类似地，较低的阈值参数值会降低节点阈值，从而降低漏报的可能性。

**初始学习周期**：ILP 是从 VAD 开启瞬间开始 VAD 算法调整自身以适应背景噪声环境所花费的时间。方程式 1 显示了如何计算 VAD\_ILP 参数。

$$\text{Initial learning period (s)} = \frac{ILP_{10}}{(256 \times 8000)} \quad (1)$$

其中

- $ILP_{10}$  是以十进制形式表示的 ILP 寄存器值

默认值 (0x001F4000) 对应于 1s。表 2-9 显示了控制 VAD\_ILP 参数的寄存器。

表 2-9. 用于初始学习周期的可编程系数寄存器

系数	页	寄存器	复位值	说明
VAD_ILP	0x07	0x7C	0x00	ILP 字节 [31:24]
	0x07	0x7D	0x1F	ILP 字节 [23:16]
	0x07	0x7E	0x40	ILP 字节 [15:8]
	0x07	0x7F	0x00	ILP 字节 [7:0]

**保持计数器**：在检测到语音活动时，VAD 算法会生成一个中断，如果该中断被编程为高电平有效，则在检测到语音时中断变为高电平 ( 逻辑 1 )，当没有语音时变为低电平 ( 逻辑 0 )。在变为低电平之前，语音活动停止后中断保持高电平的时长由保持计数器计数决定。方程式 2 显示了如何计算 VAD\_HOC 参数。

$$\text{Hold over counter (s)} = \frac{HOC_{10}}{(256 \times 8000)} \quad (2)$$

其中

- $HOC_{10}$  是以十进制形式表示的 HOC 寄存器值

默认值 (0x00032000) 对应于 100ms。表 2-10 显示了控制 VAD\_HOC 参数的寄存器。

表 2-10. 用于保持计数器的可编程系数寄存器

系数	页	寄存器	复位值	说明
VAD_HOC	0x08	0x0C	0x00	HOC 字节 [31:24]
	0x08	0x0D	0x03	HOC 字节 [23:16]
	0x08	0x0E	0x20	HOC 字节 [15:8]
	0x08	0x0F	0x00	HOC 字节 [7:0]

**唤醒等待**：如果 VAD 被编程为处于自动模式，则在检测到语音时，它将自动开启 ADC 并开始记录，同时还会检查是否存在语音活动。唤醒等待是 VAD 在进入记录模式后暂停的时长，经过该时长后会恢复。[方程式 3](#) 显示了如何计算 VAD\_WW 参数。

$$Wakeup\ Wait\ (s) = \frac{WW_{10}}{(256 \times 8000)} \quad (3)$$

其中

- $WW_{10}$  是以十进制形式表示的唤醒等待寄存器值

默认值 (0x01388000) 对应于 10s。[表 2-11](#) 显示了控制 VAD\_WW 参数的寄存器。

**表 2-11. 用于唤醒等待的可编程系数寄存器**

系数	页	寄存器	复位值	说明
VAD_WW	0x08	0x08	0x01	WW 字节 [31:24]
	0x08	0x09	0x38	WW 字节 [23:16]
	0x08	0x0A	0x80	WW 字节 [15:8]
	0x08	0x0B	0x00	WW 字节 [7:0]

**阈值**：阈值 (TH) 控制决策树节点的决策边界。较高的值将增加决策树所有节点的节点阈值，从而降低误报的可能性。类似地，较低的阈值参数值会降低节点阈值，从而降低漏报的可能性。因此，可以调整阈值参数，以在漏报和误报之间达到适当的平衡。

[方程式 4](#) 显示了如何计算 VAD\_TH 参数。

$$Threshold_{new} = Threshold_{default} \times 10^{\frac{thr}{20}} \quad (4)$$

其中

- thr 是以 dB 为单位的阈值 ( - 20dB 至 0dB )
- $Threshold_{default}$  是以十进制表示的阈值寄存器中的默认值 (16777216)

默认值 (16777216) 对应于 0dB。[表 2-12](#) 显示了控制 VAD\_TH 参数的寄存器。

**表 2-12. 用于阈值的可编程系数寄存器**

系数	页	寄存器	复位值	说明
VAD_TH	0x09	0x50	0x01	TH 字节 [31:24]
	0x09	0x51	0x00	TH 字节 [23:16]
	0x09	0x52	0x00	TH 字节 [15:8]
	0x09	0x53	0x00	TH 字节 [7:0]

### 3 VAD 结果

本节介绍 VAD 结果。算法性能由 ROC 曲线显示，该曲线说明了不同工作阈值 ( - 12dB 至 - 3dB ) 下的检测性能。以下 ROC 图显示的是来自 Aurora 噪声数据库的噪声场景 ( 图 3-1 表示汽车噪声，图 3-2 表示餐厅噪声，图 3-3 表示地铁噪声 ) 和来自 NOIZEUS 语音数据库的语音信号。测试向量是通过以所需的 12、18 和 24dB SNR ( SNR 是语音功率级别和噪声信号功率级别之间的差值 ) 混合噪声和语音信号来生成的 ( 例如，12dB SNR 意味着噪声功率级别比语音功率级别低 12dB )。工作点对于 12dB 阈值而言位于最左上角，并随着阈值的增加而向右移动，对于语音命中率和非语音命中率，图 3-4 ( - 7dB 阈值 ) 指示更高的性能。

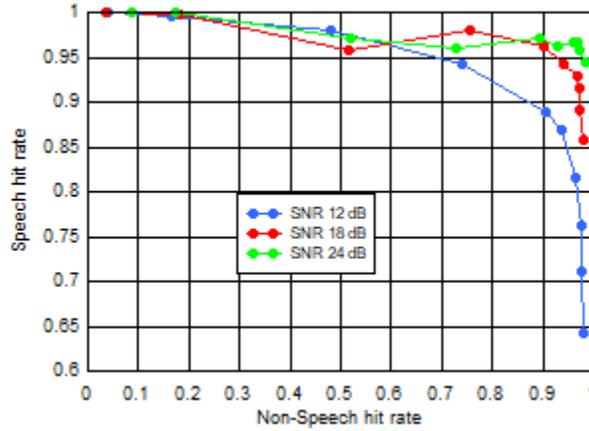


图 3-1. 针对汽车噪声的非语音命中率与语音命中率

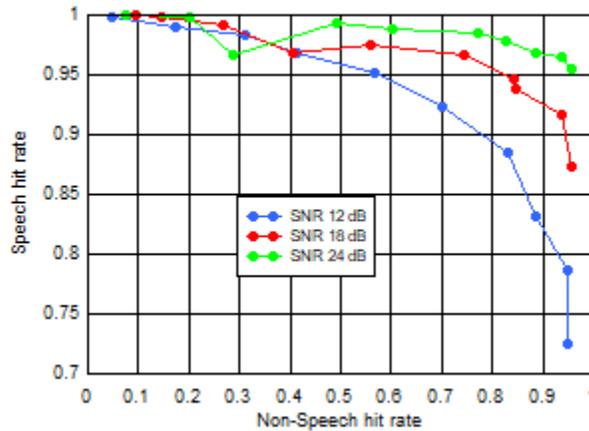


图 3-2. 针对餐厅噪声的非语音命中率与语音命中率

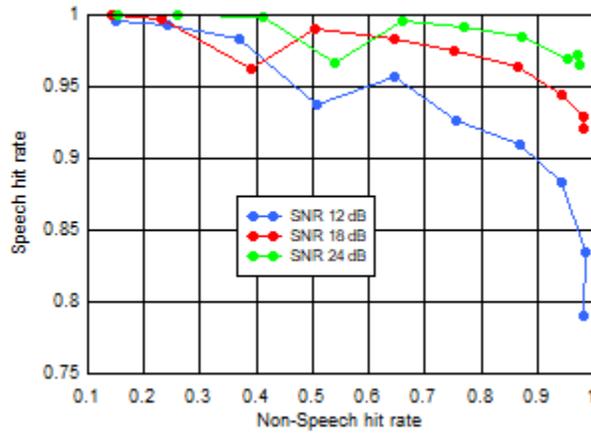


图 3-3. 针对地铁噪声的非语音命中率和语音命中率

在分析收集到的数据之后，选择了  $-7\text{dB}$  阈值，以在不同噪声类型下提供最佳语音命中率和非语音命中率。针对不同噪声类型，在  $-7\text{dB}$  阈值下的 ROC 曲线如图所示。

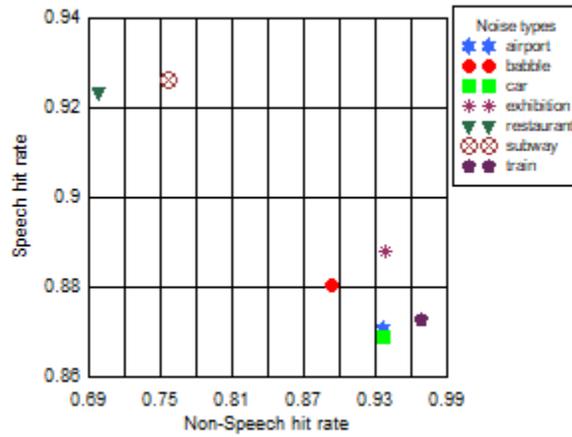


图 3-4. 针对 12dB SNR 在  $-7\text{dB}$  阈值下的非语音命中率和语音命中率

## 4 示例

本节提供了两个有关如何配置 VAD 的示例。

**示例 1：**以下示例代码显示了在通道 1 上使用 VAD 和模拟麦克风所需的配置。

```
# Key: w 9C XX YY ==> write to I2C address 0x9C, to register 0xXX, data 0xYY
#           # ==> comment delimiter
# See the corresponding EVM user guide for jumper settings and audio connections.
#####
#
#
# Power up IOVDD and AVDD power supplies
# Wait for 1ms.
#
w 9C 00 00 # Goto Page 0
w 9C 02 81 # Wake-up device by I2C write into P0_R2 # Exit Sleep mode
d 10      # Wait for 16 ms
w 9c 00 00 # go to page 9
w 9c 50 00 72 59 dc # -7dB threshold
w 9C 00 00 # Goto Page 0
w 9c 3c a0 # channel 1 set to LINE , SE , AC , 2.5k , DRE disabled
w 9c 73 80 # ip channel 1 enabled
w 9c 74 80 # ASI out ch 1 enabled
w 9c 21 00 # GPIO1 as IRQ (latched only) driving high and low
w 9c 00 01 # go to page 1
w 9c 1e 00 # Channel 1 as VAD input, User mode , internal clk
w 9c 1f 00 # VAD disabled in recording mode
w 9c 00 00 # go to page 0

w 9c 75 21 # Enable VAD
```

**示例 2：**以下示例代码显示了在通道 1 上使用 VAD 和数字麦克风所需的配置。

```
# Key: w 9C XX YY ==> write to I2C address 0x9C, to register 0xXX, data 0xYY
#           # ==> comment delimiter
# See the corresponding EVM user guide for jumper settings and audio connections.
#####
#
#
# Power up IOVDD and AVDD power supplies
# Wait for 1ms.
#
w 9C 00 00 # Goto Page 0
w 9C 02 81 # Wake-up device by I2C write into P0_R2 # Exit Sleep mode
d 10      # Wait for 16 ms
w 9c 00 00 # go to page 9
w 9c 50 00 72 59 dc # -7dB threshold
w 9C 00 00 # Goto Page 0
w 9c 3c 40 # Ch1 configured as Digital microphone PDM input
w 9c 41 40 # Ch2 configured as Digital microphone PDM input

w 9c 73 C0 # ip channel 1 and channel 2 enabled
w 9c 74 C0 # ASI out ch 1 and ch 2 enabled
w 9c 21 21 # GPIO1 as IRQ
w 9c 22 41 # GPO1 as PDM CLK
w 9c 2b 45 # gpil as Ch1/2 DMIC data
w 9c 20 00 # gpil neg edge as chl

w 9c 00 01 # go to page 1
w 9c 1e 00 # Channel 1 as VAD input, User mode , internal clk
w 9c 1f 00 # VAD disabled in recording mode
w 9c 00 00 # go to page 0w 9c 75 21 # Enable VAD
```

## 5 相关文档

请参阅如下相关文档：

- TLV320ADC6120
  - 德州仪器 (TI), [TLV320ADC6120 2 通道、768kHz、Burr-Brown™ 音频 ADC 数据表](#)
  - 德州仪器 (TI), [具有 106dB SNR 的 TLV320ADC6120 立体声通道、768kHz、Burr-Brown™ 音频 ADC 评估模块](#)
- TLV320ADC5120
  - 德州仪器 (TI), [TLV320ADC5120 2 通道、768kHz、Burr-Brown™ 音频 ADC 数据表](#)
  - 德州仪器 (TI), [具有 106dB SNR 的 TLV320ADC5120 立体声通道、768kHz、Burr-Brown™ 音频 ADC 评估模块](#)
- TLV320ADC3120
  - 德州仪器 (TI), [TLV320ADC3120 2 通道、768kHz、Burr-Brown™ 音频 ADC 数据表](#)
  - 德州仪器 (TI), [具有 106dB SNR 的 TLV320ADC3120 立体声通道、768kHz、Burr-Brown™ 音频 ADC 评估模块](#)
- 德州仪器 (TI), [PCM6120-Q1 2 通道、768kHz、Burr-Brown™ 音频 ADC 数据表](#)
- 德州仪器 (TI), [PCM5120-Q1 2 通道、768kHz、Burr-Brown™ 音频 ADC 数据表](#)
- 德州仪器 (TI), [ADCx120EVM-PDK 用户指南](#)
- 德州仪器 (TI), [PurePath™ Console](#)

## 6 修订历史记录

### Changes from Revision \* (December 2021) to Revision A (April 2022)

Page

- |                                    |   |
|------------------------------------|---|
| • 添加了 PCM5120-Q1 和 PCM6120-Q1..... | 1 |
|------------------------------------|---|

## 重要声明和免责声明

TI“按原样”提供技术和可靠性数据（包括数据表）、设计资源（包括参考设计）、应用或其他设计建议、网络工具、安全信息和其他资源，不保证没有瑕疵且不做任何明示或暗示的担保，包括但不限于对适销性、某特定用途方面的适用性或不侵犯任何第三方知识产权的暗示担保。

这些资源可供使用 TI 产品进行设计的熟练开发人员使用。您将自行承担以下全部责任：(1) 针对您的应用选择合适的 TI 产品，(2) 设计、验证并测试您的应用，(3) 确保您的应用满足相应标准以及任何其他功能安全、信息安全、监管或其他要求。

这些资源如有变更，恕不另行通知。TI 授权您仅可将这些资源用于研发本资源所述的 TI 产品的应用。严禁对这些资源进行其他复制或展示。您无权使用任何其他 TI 知识产权或任何第三方知识产权。您应全额赔偿因在这些资源的使用中对 TI 及其代表造成的任何索赔、损害、成本、损失和债务，TI 对此概不负责。

TI 提供的产品受 [TI 的销售条款](#) 或 [ti.com](#) 上其他适用条款/TI 产品随附的其他适用条款的约束。TI 提供这些资源并不会扩展或以其他方式更改 TI 针对 TI 产品发布的适用的担保或担保免责声明。

TI 反对并拒绝您可能提出的任何其他或不同的条款。

邮寄地址：Texas Instruments, Post Office Box 655303, Dallas, Texas 75265

Copyright © 2022，德州仪器 (TI) 公司