REPORT
January 2015
updated July 2015

‘LOLA in a BOX’

Dr Tania Lisboa, Videoconference Programme Manager, Royal College of Music

Matt Parkin, Head of Studios, Royal College of Music

LOLA has been used primary for music rehearsal as well as small group coaching so far. Performances have also used the system and quality seems to be improving constantly. The system does not require a highly complex audio setup, as no audience is involved and at the Royal College of Music (RCM), we would like musicians to be as free as possible to schedule LOLA sessions according to their needs and existing contacts internationally. At the moment, they are relying on intensive technical support from the RCM studios and therefore, we have been searching for alternative ways of making LOLA more easily accessible and hopefully set-up in a way that requires less technical support. We envisaged placing the system in a specially allocated practice room where it would be set up for self operation without specialist support. This would help maximise the use of the system and it would allow musicians to build more partnerships across a much broader spectrum than currently practiced internationally. This could lead to increased and more performances as well as more creative uses of LOLA from both students and staff**.**

Aims and scope of the project

This project was set out to discover whether a simplified hardware configuration could be devised to accommodate the majority of small rehearsal/coaching sessions and allow self-operation of the system by musicians and teachers. A key element of the project was expected to be the choice and physical positioning of audio and video hardware in order to make the system as straightforward as possible for the broadest range of musicians, with minimal user adjustment. The project assumed that expert technical input, probably on-site, would be required to build, configure and maintain the system but it should not be required for day-to-day use by small groups. Public performances and large groups (greater than 2 or 3 players at either end) are out of scope for this project. We have also considered changing the title of this project to ‘*Self-op LOLA system’* as this will make it easier to publicise (and explain) the new set-up and user possibilities within the Institution.

Equipment

The development and design of the *Self-op LOLA system* commenced in early 2014. From the outset it was recognised that the system design would largely be determined by the current state of development of the LOLA software and its supported hardware options.

The early stages of the design development led us to identifying the following principles for the system design:

Self-Op LOLA: initial design principles in order of priority

1. Ease of use – must be self operable
2. Excellent audio monitoring, in optimum position
3. Portability – able to move system fairly easily to a different room
4. High quality audio capture
5. Very high speed monitor (low refresh rate)
6. Video camera positioned directly above or below monitor
7. Good video delivery
8. Recording function if possible

The first step was to choose and build a suitable LOLA PC, with the necessary camera and audio/video hardware. When the project was initially proposed, it was expected that the design would be based on the hardware used in all LOLA systems to date, and the project was scoped and costed on that basis. However, at the beginning of this project, the LOLA team at *Conservatorio G. Tartini* announced the forthcoming release of version 1.4 software. Significantly, this promised higher definition video resolutions for the first time, with support for high resolution digital cameras and framegrabbers via coaxpress[1] and USB3 connections. Previous versions had always required analogue cameras, which are now mostly obsolete and increasingly difficult to source.

In principle a USB3 camera seemed like the ideal option for our project, with a simpler interface, lower cost and even the potential to run the system on a laptop. However by late Spring, the capabilities of a USB3 camera were still uncertain and the beta releases of LOLA version 1.4 were principally proven with coaxpress cameras. Therefore, we decided to proceed with version 1.4 and source a coaxpress camera, which offered a number of advantages over legacy analogue and also future USB3 cameras, such as:

1. Higher resolution

[1] http://www.coaxpress.com/coaxpress.php

2. Single BNC cable for power and data
3. Capable of long (up to 100m) cable runs between PC and camera
4. Reasonably wide choice of lenses available
5. Cameras readily available and supported by manufacturers

The model chosen was the **Imperx Bobcat B1411C camera**[2], a C-mount coaxpress camera with ½" sensor and native resolution of 1392x1040 pixels. This choice came at an increased cost to the project budget and procurement of the camera, which is still a specialist item, took several weeks and led to delays in the testing and designing development stage proposed. It was interesting to know that the RCM was the first site to use this precise model of camera with LOLA.

Having settled on a camera choice, PC selection and configuration was undertaken by the project team as well as the RCM ICT department. ICT, studios and project team then worked closely on a detailed design, build and testing of the system.

In terms of equipment, we initially hoped to procure as compact a PC as possible, in order to make the system reasonably portable or to be able to install it unobtrusively and with minimal impact in a small/medium practice room. However, evaluation the detailed hardware requirements for LOLA version 1.4 it rapidly became apparent that a full size tower PC would still be required for optimum reliability and compatibility, especially to allow for the full height PCIe coaxpress frame grabber and audio cards. In consultation between the project team, RCM ICT and external suppliers, we agreed to choose a fairly high specification PC, settling on a customised version of the **Hewlett Packard Z420**[3]. It was felt this would ensure full compatibility with the still-developing requirements for LOLA version 1.4 and the most reliable operation, especially if higher definition operation was to be possible.

The full specification of the PC eventually built for this project is shown in Appendix 1.

Further hardware choices made initially were as follows:

Frame grabber	BitFlow Karbon-CXP2
Graphics card	Asus Nvidia GeForce GTX 650
Audio card	RME HDSPe-AIO
Fast-refresh (gaming) full-HD LCD 27" PC monitor	Asus
Microphones (for testing*)	Schoeps, Neumann, Sony etc.
Loudspeakers (for testing*)	Genelec 1032 active
Audio mixer (for testing*)	Mackie 1402-VLZ3 analogue

*Access to RCM Studios stock allowed a wide range of choices for microphones, speakers and audio mixers to be used during testing and development of the concept and design, though it was anticipated that the final assembled system would use specially sourced equipment to suit the final design.

[2] http://imperx.com/ccd-cameras/b1411/
[3] http://www8.hp.com/uk/en/products/workstations/product-detail.html?oid=5225033#!tab=features

Having assembled the PC and various components, a great deal of technical testing was undertaken with remote sites including the Royal Danish Academy of Music in Copenhagen and LOLA HQ in Trieste. Initially we encountered significant problems getting the camera to work reliably and it soon became apparent that the B1411C camera did not work as expected. Although the B1410C was believed to behave identically to the more expensive B1410C camera used by the LOLA team in testing, our implementation revealed undocumented differences. This eventually required specific camera files to be developed by the LOLA developers, using our system as a remote test node, for the B1411C to operate successfully on LOLA at all the necessary resolutions. This process took around a month but resulted in a fully working, high resolution LOLA node which performed excellently in connectivity tests.

Design

Initial research and considerations

The RCM's prior experience with LOLA provided a strong starting point for the design. Initial findings included the following:

1. Microphone and speaker placement must primarily be determined by the need to minimise feedback, as there is no echo cancellation in the LOLA software

2. Most musicians quickly grasp the importance and impact of loudspeaker positioning, since it immediately affects their own experience. However they are unlikely to understand or consider the importance of microphone positioning – which nonetheless has a critical impact on the remote musician's experience of using LOLA.

3. Fairly well-damped room acoustics tend to be preferable for the above reason, but this should be balanced by the need for the musicians to have a comfortable acoustic playing environment – so for classical music, a totally 'dead' room is unlikely to be suitable.

4. The main picture from the remote site should be as large as possible on the screen – probably full screen – though musicians do also like the reassurance of seeing a small image of their own transmitted image from the local camera.

5. Excellent eye contact between musicians is important in terms of building rapport (especially between musicians who have never played together previously or, as in many cases, never even met before). Therefore screens and cameras should be positioned as close together as possible, near to the performers and at eye height.

6. One of the most obvious 'fault conditions' of LOLA in use is audio clipping – which generates loud digital clicks. However due to the excellent audio quality of the link, it is generally possible to maintain a high level of audio headroom in order to keep clear of 0dBFS clipping level. Any self-op system should therefore be set at low enough gain to avoid clipping levels, and/or offer simple and immediate control of mic gain so musicians can quickly adjust levels to avoid the clipping.

Many of the initial test sessions were conducted with Edinburgh Napier university. Gill Davies, who led the Napier end of that research, provided the following useful summary of some initial findings regarding camera and screen positioning and the general environment:

> *From the first two sessions (clarinet and piano) and (guitar/piano and piano/cello), the musicians involved all said that positioning of the screen was important to ensure they had eye contact, if possible, with the remote musician. [We should therefore consider] telling the musicians how to get the best camera angle – it could be worth them picturing the view they would like first and then framing the shot accordingly.*
>
> *Another musician said that it was more difficult to establish a rapport with the remote musician because there were so many other people in the room (we were in a recital room at Napier and there were a few folk hanging around).*
>
> *In terms of operating LOLA themselves – most musicians said they would not be comfortable with setting up the LOLA node. One musician said he would be comfortable operating the system himself (inputting the remote IP address and connecting).*
>
> *One musician said that training in the system should include a session on how to get the system up and running if it crashes.*

Developing the system design

The design was based on the following consideration::

1. Screen and camera choice, mounting and positioning – these are design factors and once decided, the musicians should not need to modify either position or make any adjustments. These choices are more or less independent of any decisions about audio equipment.

2. Loudspeaker positioning – this is a critical issue and musicians have strong opinions about this.

3. Microphone positioning – musicians are unlikely to have opinions about this, but the choice of microphone and how it is mounted/positioned depends entirely on the speaker positioning, therefore cannot be decided prior to settling on the best solution for item 2.

4. Control interfaces – providing the simplest possible software and/or hardware control interfaces to enable musicians to operate the system on their own with minimal instruction.

5. Assembly of a fully integrated layout and physical casing/mounting design to accommodate the above factors, resulting in a finished design and prototype physical product.

1. Camera/screen positioning

From the outset and based on previous LOLA experiences, it was decided to try and locate the camera directly above and in the centre of the monitor. A **27" high speed LCD panel** was selected to give a good balance of size vs portability and allow the camera to remain at eye height while being positioned directly on top of the screen.

A number of tests were undertaken which revealed the importance of lens size: at lower definitions LOLA crops the image, using just the inner pixels of the sensor, it was found that lower resolutions required wider lenses.

The lense initially purchased with the camera was a 12mm F1.4 lens with manual iris, which provided a reasonable image for medium size ensembles at the highest resolution. However, during testing, it was found that switching to lower resolution was sometimes beneficial, for example to cope with bandwidth restrictions when network issues could not be resolved.

Attempts were made to identify a suitable C-mount zoom lens but a cost effective option was not readily available, so additional wider lenses were purchased instead to provide flexibility for working at different resolutions and with different types of ensembles.

Ideally, a motorised zoom lens would probably be the optimum solution – perhaps even controlled directly by the LOLA software in order to allow different resolutions with the same effective image size. However this is beyond the budget and capabilities of the current project.

Another idea suggested during testing: in order to maximise eye contact, a 'teleprompter style' reflective glass screen and hood would allow the camera to be mounted directly behind the image. This concept requires testing and development, which we hope will take place prior to the presentation of this project in May 2015.

2. Loudspeaker positioning

The starting point for speaker position was to localise the remote sound with the remote image – i.e. to place the speaker(s) as close to the screen as possible. However, during previous sessions we had found it preferable to place the speakers closer to the players, for example to their side or just behind them. Therefore, we experimented with both options in the testing phase of this project.

Tests with musical duos of various kinds suggested that co-locating speakers with the screen was often the preferred option. Even when other options were preferred, there was little difference.

Testing with a group of recorders (up to 4 players at RCM joining one further player at the Royal Danish Academy of Music (RDAM) allowed us to further experiment the 'co-location' principle. During the course of a lengthy session with different sizes of ensemble at our end we tried to optimise the naturalistic experience for players at both ends.

During these tests it became apparent that stereo image was also a significant factor in helping musicians feel comfortable playing with more than one remote partner, affecting both speaker and microphone placement. With more than one player each end, the ability to determine convincing stereo localisation of each remote instrument (especially similar or identical instruments) was a critical factor in following each other and playing reliably together in time.

The pictures below demonstrate the progression of positions attempted during this session:

(a) loudspeakers directly in front of the screen, to left and right

(b) loudspeakers closer to the musicians to reduce feedback and improve stereo image

(c) loudspeakers angled outwards for additional stereo separation and to point more away from the stereo mic

As a result of the various tests it was concluded that for straightforward rehearsal and coaching purposes, the optimum loudspeaker mounting would be just in front of the screen and positioned to maximise good stereo separation. Additionally the speakers would be good quality monitor loudspeakers suitable of providing a realistic level of replay for classical instruments without noticeable distortion, and ideally fixed permanently to avoid the need for user positioning. A simple volume control should be provided for controlling the loudspeaker level.

3. Microphone positioning

Having decided the general requirements for loudspeaker positioning, it was possible to then review the options for mic placement. During testing before and during the period of this project, a wide range of typical 'classical recording' microphone placements have been used. These have also had to accommodate the additional critical requirement of avoiding feedback. It is also important to obtain clear and natural speech pickup, to make the rehearsal and coaching experience as normal as possible.

Stereo image has also been shown to be an important factor and as a result, various combinations of stereo mic pair have been trialled and used, especially for sessions with more than one player at each end.

For larger scale, engineered sessions, we have found it is useful to have the following options available, under the control of a trained operator:

- Instrument microphone(s) positioned reasonably close to the player/singer, either mono or stereo depending on the instruments. These would generally still be no closer than about 50-100cm away from the instrument in order to achieve a natural sound – i.e. further away than might be the norm for live sound amplification or pop music recording.
- Slightly more distant stereo pickup to make the sound more natural – especially for those instruments where early reflections and reverberation from the room help aid the perception and judgement of the quality of the playing/singing

- Separate speech microphones, e.g. a hypercardioid or gun microphone or wireless lapel system, especially when working with loud instruments where the speech levels are relatively quiet. These are not often necessary – but do require active control/operation by an operator if used

Based on the above experience, and the scaled down requirements for the present project, we identified that a single stereo pair would be suitable, ideally positioned slightly in front of and in between the speakers – similar to the stereo mic position shown in the above recorder quartet sessions. Because public performances and larger groups are out of scope for this project and in order to provide a simple, compact and user-friendly system for small scale collaborations, we decided the room and speech mics can be disregarded for the proposed system.

Based on the above the original preferred microphone option was the **Sony ECM-680S**[4] stereo microphone which provides a stereo L/R output via internal decoding of the mid-side (MS) capsule outputs and provides a relatively high degree of directivity and simplicity of operation for non-technical personnel. Subsequently due to the phantom powering capabilities of the Allen & Heath mixer, it was necessary to identify an alternative microphone capable of being powered by a 15V supply. The **Audio Technica AT8022** was identified and has proven to be well suited to the task.

4. Control interfaces

From the outset it was identified that a crucial aspect of the project's success would be to simplify the setup and control of the system to make it possible for non-technical personnel to use it with minimal instruction. We considered a number of software enhancements and add-ons to simplify the operation of the LOLA software itself and the required audio/video drivers. Once set up in a stable configuration, the actual activation and use of the LOLA software to undertake sessions is remarkably straightforward. The frequent updates to the LOLA GUI are already addressing many of the issues and constraints identified in earlier versions, such as the ability to easily resize video windows or diagnose basic connectivity problems. It was further recognised that extensive software development was likely to be beyond the capabilities and resources of the present project. Therefore, we decided to concentrate our resources on making the external hardware control and setup as simple as possible, to match the ease of use of the software. Video setup issues are almost entirely a matter of physical layout and configuration of the screen and camera, as described above. However, audio levels are still bound to require some manual intervention and a conventional professional audio mixer was thought to too complicated for the purpose, with additional potential for settings to be incorrectly adjusted.

[4] http://www.sony.co.uk/pro/product/broadcast-products-pro-audio-wired-shotgun/ecm-680s/overview/

Following research of a wide range of analogue and digital compact audio mixers, controller, preamps and so on, we were eventually led towards a small rackmount mixer designed for commercial public address system installations, the Allen & Heath GR05[5].

This mixer combines a very simplified front panel control interface with a highly flexible series of internal configuration options, all in the analogue domain which simplifies operation and minimises latency. Additionally the mixer provides sufficient busses to control both the input and output paths of the Self Op LOLA system and, unlike many similar products, is designed for stereo operation. Although the mixer does supply microphone power it is not sufficient to power the Sony condenser microphone and therefore, an additional phantom power adapter is required.

Using the above system, the final audio control could be reduced to in effect two controls – one for the 'send' (microphone) level and another for the 'receive' (speaker) level.

5. Final design

Following selection of the above components, the final design and assembly of the prototype system was completed during the first half of 2015.

In summary, the prototype system comprises:

- A high resolution LOLA node built around the HP Z420 PC, with Imperx Bobcat B1411C coaxpress camera, Genelec professional loudspeaker monitors, Asus low latency LCD panel display and Audio Technica AT8022 stereo microphone
- Front panel control of key audio levels via simple 'volume knob' style controls
- All components fitted in a bespoke case in fixed positions to maximise usability in a practice room environment and minimise audio feedback etc.
- Simple printed and on-screen guidance for users to enable them to start and run the software and adjust the necessary controls themselves

[5] http://www.allen-heath.com/ahproducts/gr05/

6. Photographs of the finished LOLA system

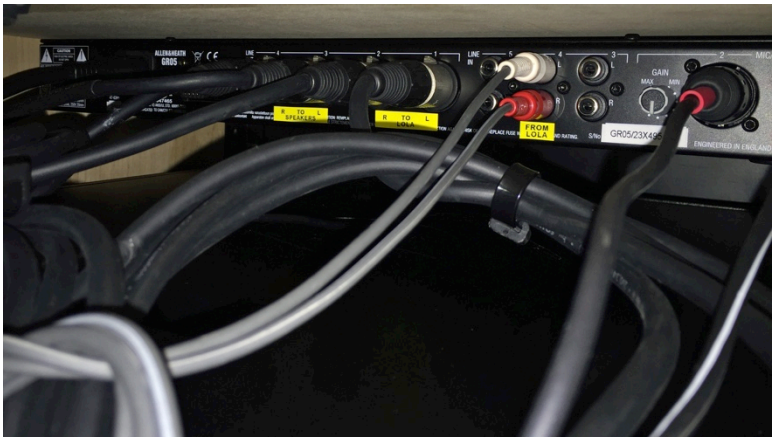
(a) The fully assembled ‘LOLA in a Box’ apparatus

(b) Positioning of microphone, audio controls and PC keyboard

(c) Simplified audio controls with two knobs for student operation

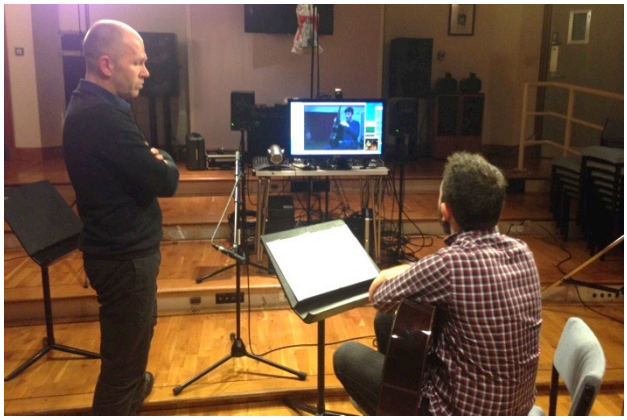
(d) Camera position directly above centre of high speed HD LCD panel

(e) Rear panel labelling allowing quick and simple connection



Guitar Duets: A Case Study

After several tests mentioned above, the RCM has run a performance and coaching session with the Royal Danish Academy of Music (RDAM) to further study the implications of the above mentioned points. During this session, one RCM postgraduate guitarist played a selected duet piece consisting of two contrasting movements, with two different students from the RDAM – each performing and leading one movement of the musical piece. The RCM also had a professor present, who engaged in both teaching the group and later providing his own views of LOLA for both teaching and performing. We are very interested in the views of users and below are some of the points highlighted by the professor.



1. **Sound**

The professor had no experience of using such technology and he found that he had to 'adjust his hearing' to the sound coming from the speakers. He asked for the volume to be adjusted quite a few times during the session:

> *"To begin with, some time was needed to adjust the sound quality, position of microphones, volume levels etc. and we noticed that the natural acoustics of the studios were quite different. However, once that had been taken into account; the outcome depended more on the players´ own intensity of interpretation, meaning that a considerably live-like audio ambience had been achieved. The miniscule delay was not noticeable".*

2. **Image**

It was interesting to see that the professor suggested adjusting the sitting position of the RCM student to face the screen sideways as he would do in a public performance (rather than sitting looking straight into the screen as the team had set-up):

> *"The video aspects were somewhat more complicated. When they started playing it became clear that better visual contact was needed, something which is of critical*

importance in ensemble work. The screen at the RCM studio had been placed right in front of the player, as were the speakers, which did not represent a real-life duo situation. By changing the position of the players against the cameras their visual contact improved and so did their playing. Time was rather short, but as the session progressed the students noted less distraction from the unconventional settings and their attention could increasingly be pointed towards their playing and the music itself."

3. Quality of connection

The professor and students questioned the effect of a possible delay in transmitting in relation to *tempo* in performance. However, we suspect that this question may be due to the lack of experience of players in using the technology and in focusing on what was happening in the room (or perhaps even in the mere fact that this was the first time they performed together and without a rehearsal).

"Afterwards, one of the issues we discussed with our Danish colleagues was the fact that the duos´ tempos tended to slow down towards the end of each movement and that this might be due to the delay in transmission. Although the delay was very slight (20ms in total) the players might constantly have to be adjusting to it without being consciously aware of it. On the other hand, this might only have been due to the unusual setting, which of course needs some time to adjust to. Also, the duet was relatively new to the performers so they were quite dependent on reading their scores, resulting in less freedom in inter-performer contact."

4. Overall impression

The professor's overall view was that LOLA is a powerful tool for teaching an performing and that he would like to have more access to this type of sessions:

"LOLA is obviously a powerful and potential tool for teaching and ensemble playing. This was the first time I had attended a teaching session over high-speed internet and I was surprised to see that the technical aspects did not seem to be causing any real problems. The audio and video quality were excellent. In my opinion, areas of improvement might therefore be the physical setup itself: the position of the players, the speakers and the screens. In this sense, the RCM would definitely benefit from a larger screen. One might also point out that the RDAM studio allows for the equipment to be set in a fixed manner, in contrast to the RCM´s which has to be set up especially for each session which is more time-consuming and might therefore lead to somewhat less flexibility in experimentation. I am most grateful for having been allowed to take part in this highly interesting work and hope to be able to assist again in the future."

An international demonstration to the Institutions

A performance session to involve a larger number of staff and students was also set-up in December 2014. This session brought together members of all departments at the RCM and the RDAM for a Festival celebration of Christmas. The RCM choir, conducted by the Artistic Director Stephen Johns experienced for the first time the technological advancements in the performance world through what was seen as a 'fascinating' and 'fun session' - *A Christmas Carol presentation* in which singers in London were accompanied by organists in Copenhagen.

Although this involved a larger number of participants, the session was very successful. It was also the first time that a conductor was involved in a LOLA event and we were able to further test its reliability by analysing the gestures of a conductor working with a musician at the other location. We were concerned that the leading gestures would arrive too late for the accompanist to follow the choir but this was not the case and the session went smoothly planned without any anticipated musical problems. Members of the JANET Network were present and we believe no further information in this report is necessary.

Public presentation of project

A public presentation of the whole project including its results took place at the *Network for the Performing Arts Workshop* in May 2015. This provided an opportunity to disseminate the system as well as raising more interest within other international Institutions, which could in turn, also have an impact on connectivity at a more general level (i.e. more institutions may become interested and invest on having an easier to use LOLA system). We envisage further discussions on how LOLA could be used on a daily basis by non-technical users as a core part of the learning, teaching and performing in music. We also hope to discuss the creation

of possible online resources where online documentation about synchronous distance learning could be shared internationally. Thus, the impact of this project will go beyond designing and building ‘LOLA in a Box’ as it will provide real opportunities for users to explore its full potential musically speaking.

Further Research

We are currently discussing how members of the team could apply for funding to develop a large scale research project, based on longitudinal case studies of a chamber music groups rehearsing via LOLA. We envisage an investigation concentrating on specific performance issues directly relevant to the synchronous distance learning.

APPENDIX 1

LOLA system PC spec for this project

Hewlett-Packard Z420 Workstation (HP Z420 600W 90 Efficient Chassis)
Microsoft Windows 7 Professional Edition 64bit OS
Intel Xeon E5-1620v2 3.70Ghz 10MB 1866 4C CPU
8GB DDR3-1866 ECC (1x8GB) Unbuffered RAM
500GB 7200 RPM SATA Hard Drive
16x SuperMulti DVDRW SATA

Graphics card

Asus Nvidia GeForce GTX 650 (2GB GDDR5, PCI Expresss 3.0, HDMI, DVI-I, VGA)