An open-access database of lead-I ECG signals with practical variability scenarios

Juan Carlos Bernal-Romero
Juan Manuel Ramirez-Cortes

National Institute of Astrophysics, Optics and Electronics (INAOE)

Table of contents

Introduction

- Increase in biometric system design and implementation due to IoT, cloud computing, and intelligent systems.
- Revolution in access control systems through pattern recognition (e.g., facial recognition for unlocking phones).
- Public traits: No consent needed (e.g., voice, face on social media).
- **Non-public traits:** Require explicit consent (e.g., cerebral, cardiovascular signals).
- Unique ECG signals due to individual heart anatomy/physiology.
- Ambulatory ECG systems improve user comfort and acceptance.
- *Lead-I electrode* configuration is common for ease of acquisition.

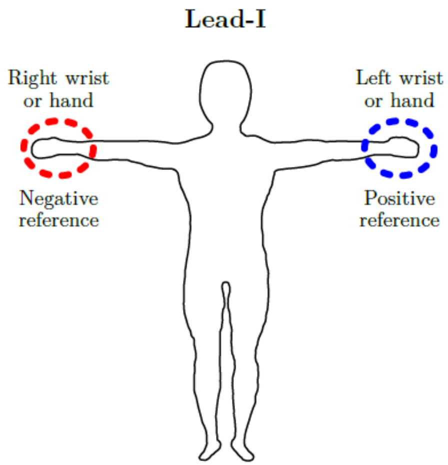


Fig. 1. Configuration Lead-I of electrodes.

Introduction

Challenges in ECG biometry systems:

- Short acquisition times (<5 seconds).
- Performance with large databases (>500 persons).
- ECG signal variability impacts recognition rates.
- High cost and computational demands for processing.

Database Contribution:

- Large, openly available lead-I ECG signal database.
- Includes various scenarios: health, activities, stress, drugs, anatomy changes, habits.

Impact:

- Supports development of robust, efficient, secure biometric systems.
- Advances recognition and diagnosis using ECG signals.

Introduction

- Decision-making based on hard and soft biometric information for more adaptive and robust systems.

Acquisition

Filter

Amplitude normalization

R-peak detection

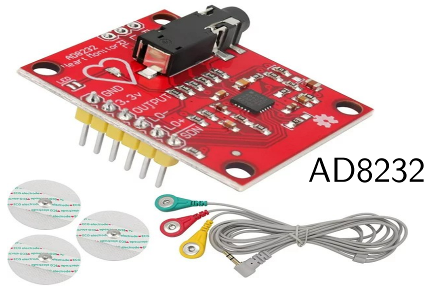
PQRST complex detection

Average PQRST complex

Heart rate

Feature extraction

Fiducial point detection



AD8232

Zero-phase filtering based on an IIR filter

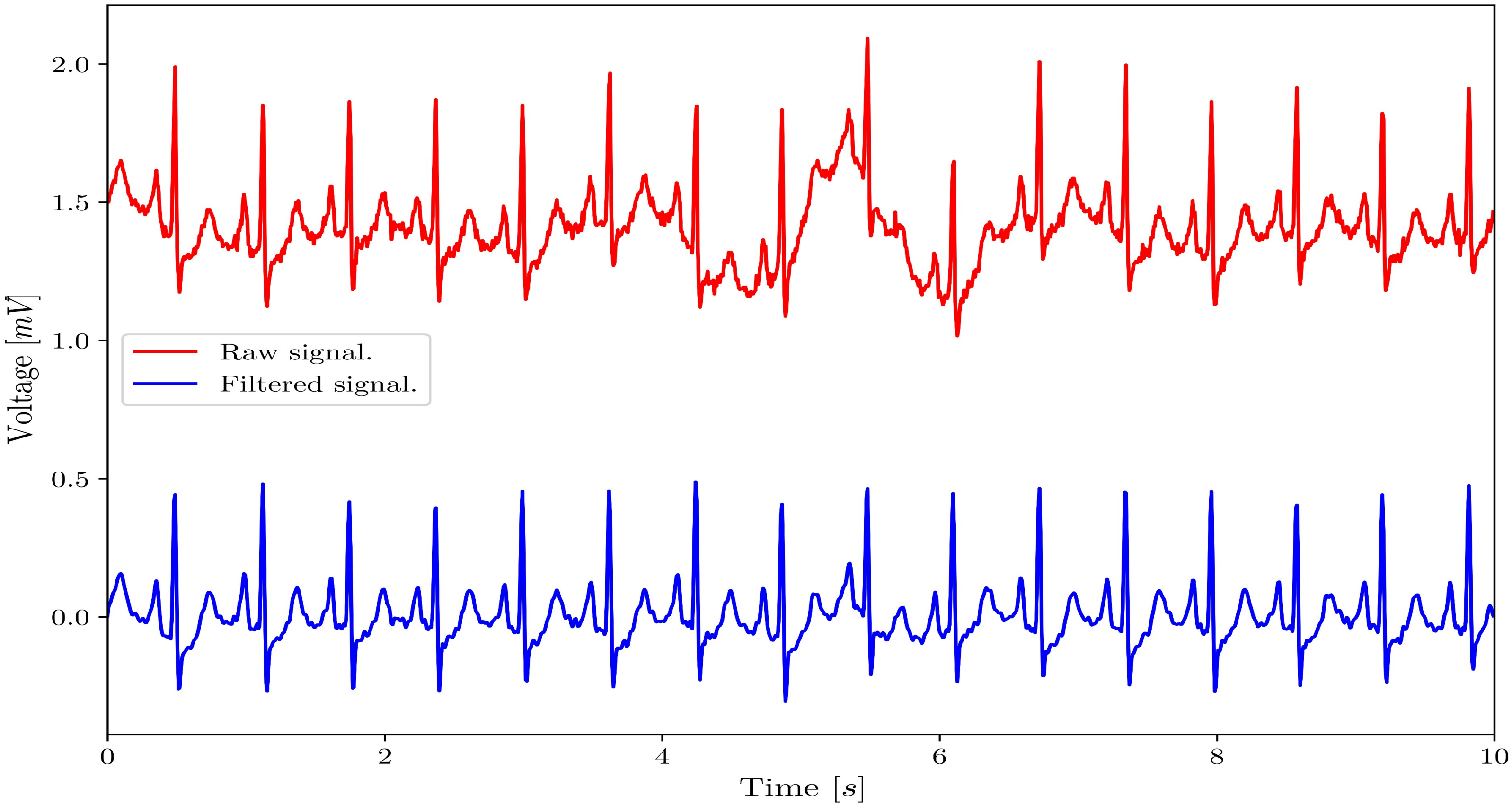


Fig. 2. Example of filtering for subject 345 (observation 2).

Zero-phase filtering based on an IIR filter

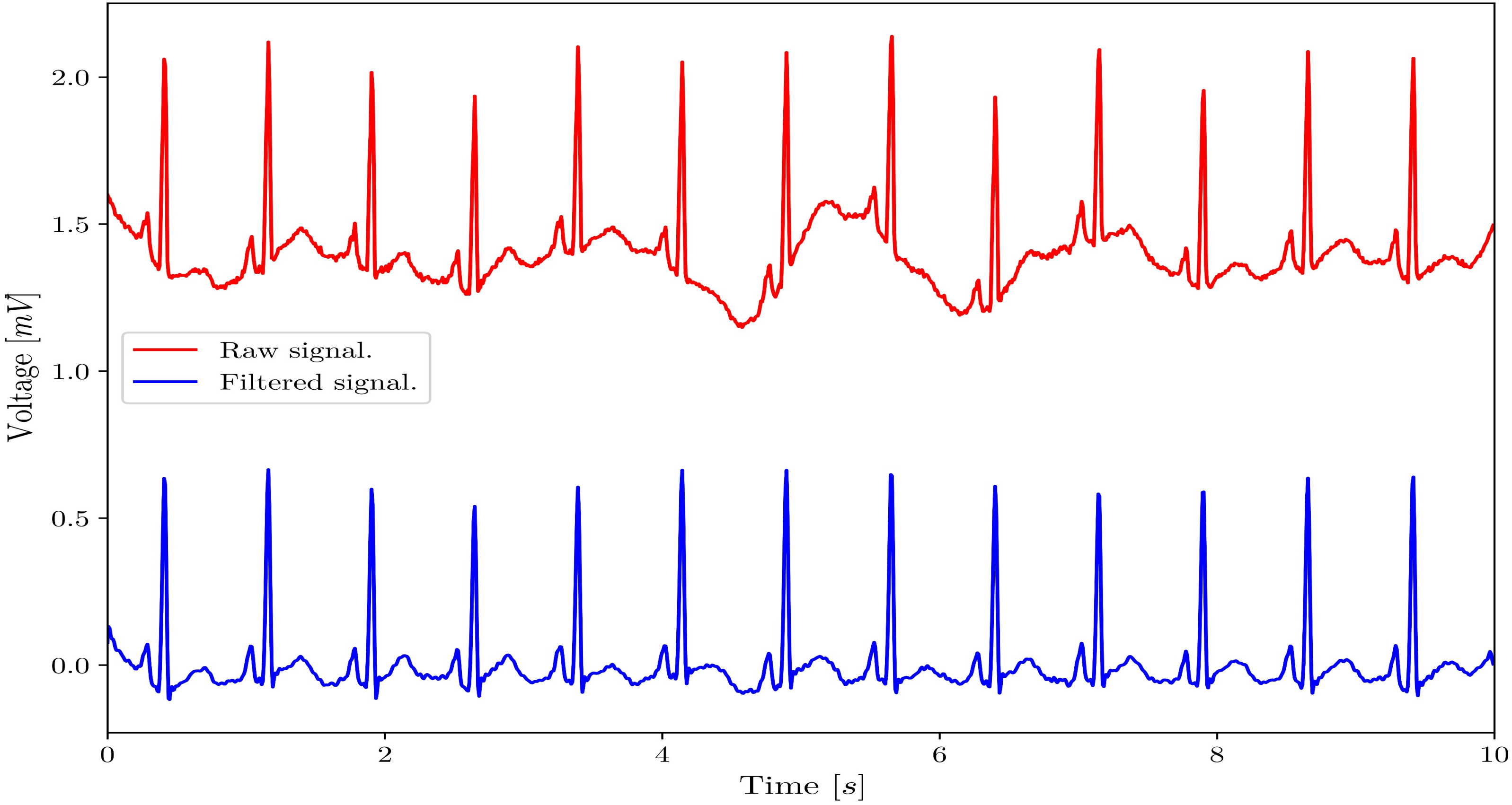


Fig. 3. Example of filtering for subject 364 (observation 4).

Normalization and PQRST complex detection

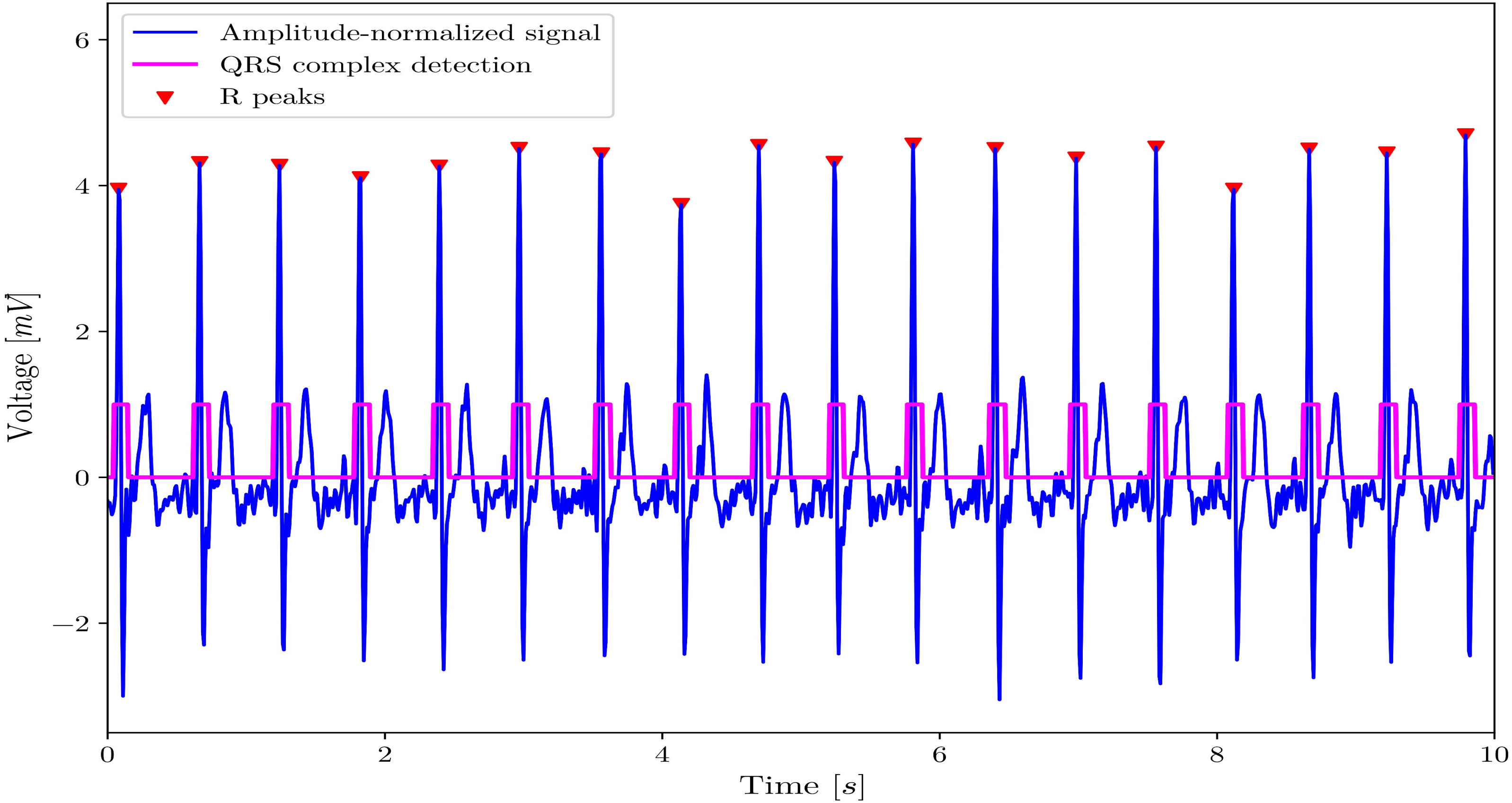


Fig. 4. Example of PQRST complex detection for subject 122 (observation 3).

Average PQRST complex and fiducial points

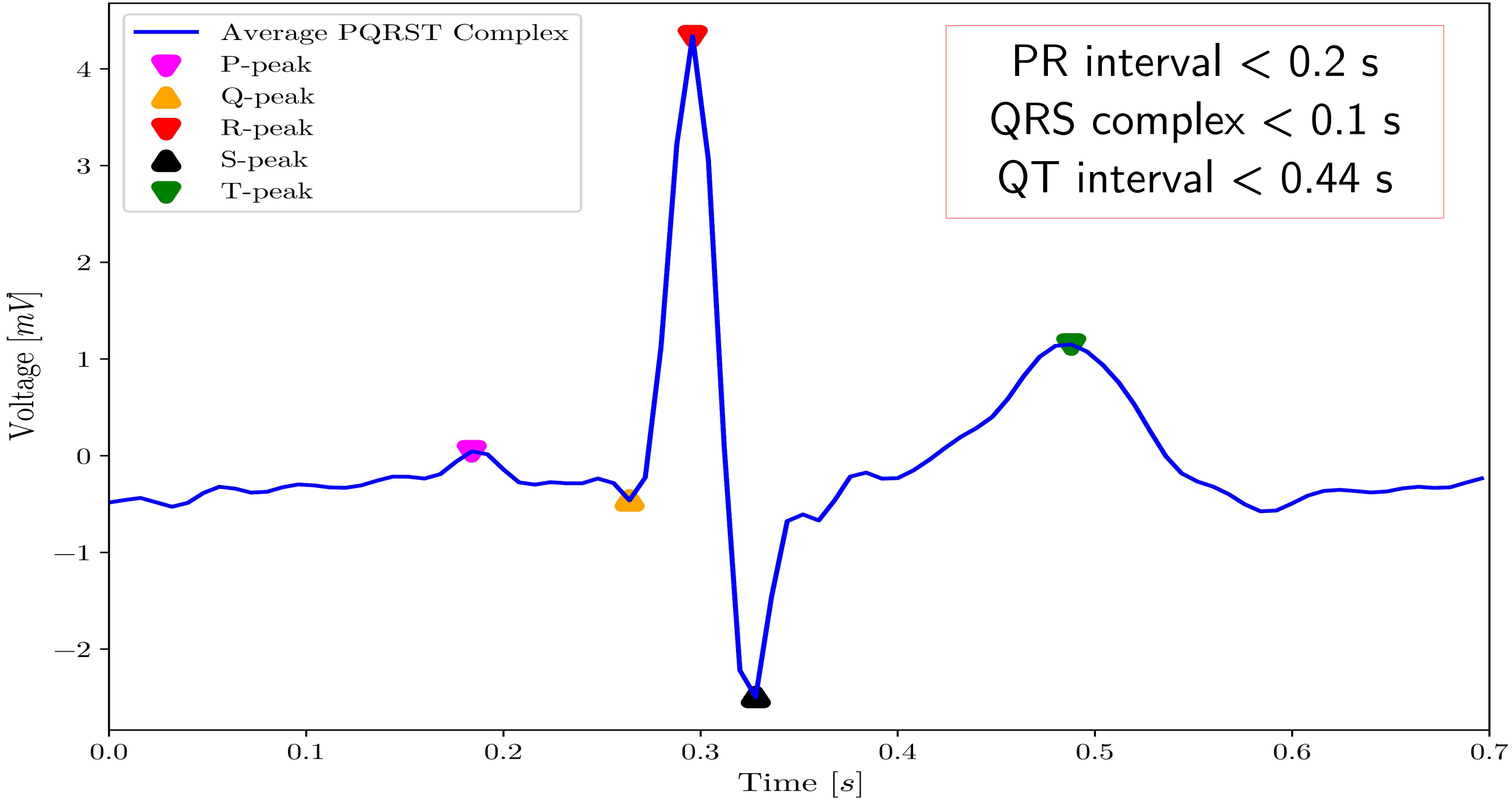


Fig. 5. Example of average PQRST complex for subject 122 (observation 3).

Scenarios of ECG signal variability

Types of ECG Signal Variability:

1. Inter-User Variability:
   - Differences in ECG signals between different individuals.
   - Essential for biometric recognition systems.
2. Intra-User Variability:
   - Differences in ECG signals from the same individual in different scenarios and times.
   - Major challenge for recognition and diagnostic systems.
   - Includes **short-term** and **long-term** changes.

Scenarios of ECG signal variability

Heart Rate Variability (HRV):

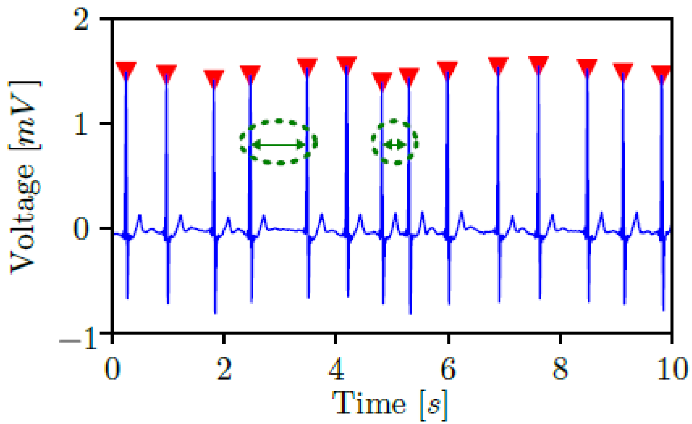


Fig. 6. Heart rate variability for a lead-I ECG recording.

- HRV refers to the variation in time intervals between consecutive heartbeats (RR intervals).
- Indicator of heart health and function.
- Short-term and long-term HRV changes provide insights into physiological and psychological states.
- The autonomic nervous system (ANS) manages HRV through the balance of the *parasympathetic* and *sympathetic* systems, ensuring homeostasis.

Heart rate variability (HRV)

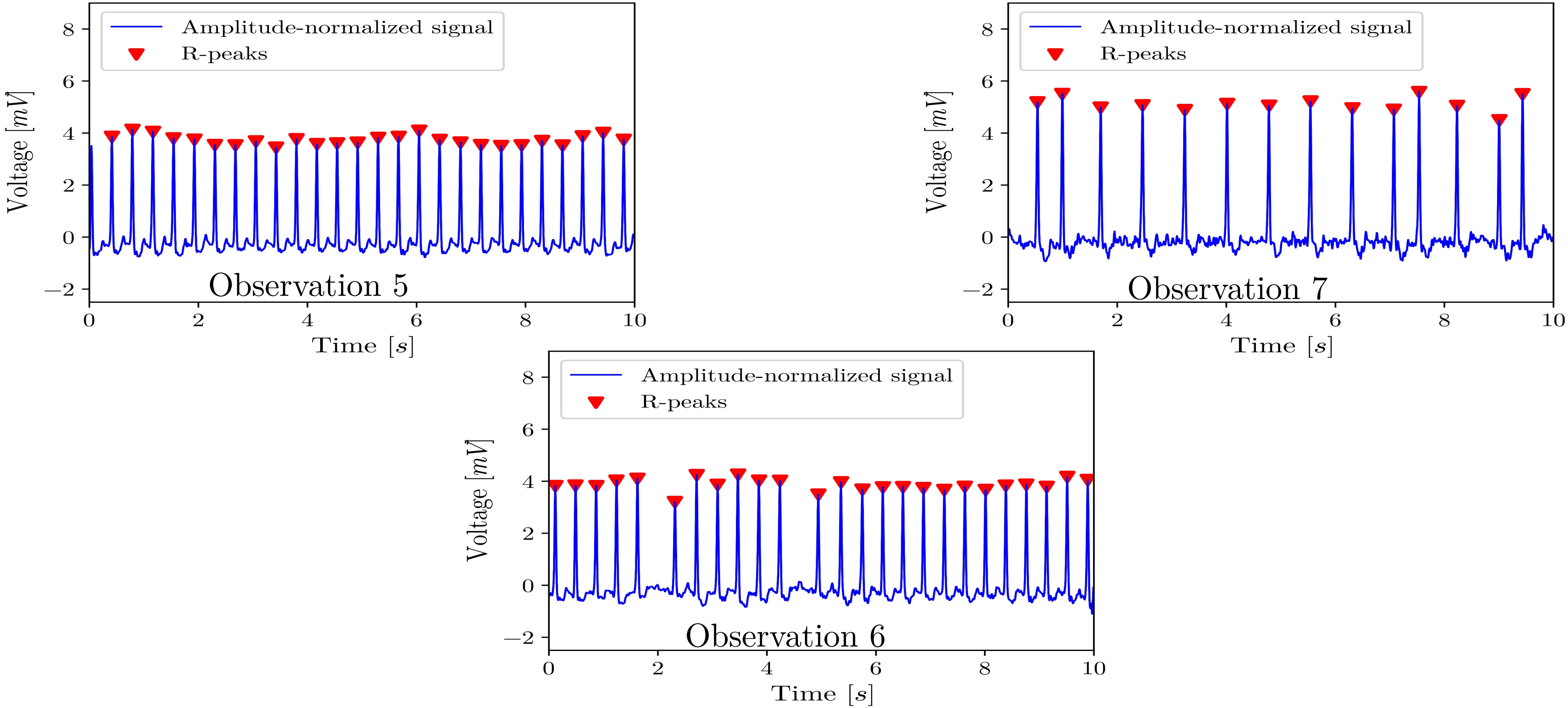


Fig. 7. Example of HRV in different observations (scenarios) for subject 414.

Heart rate variability (HRV)

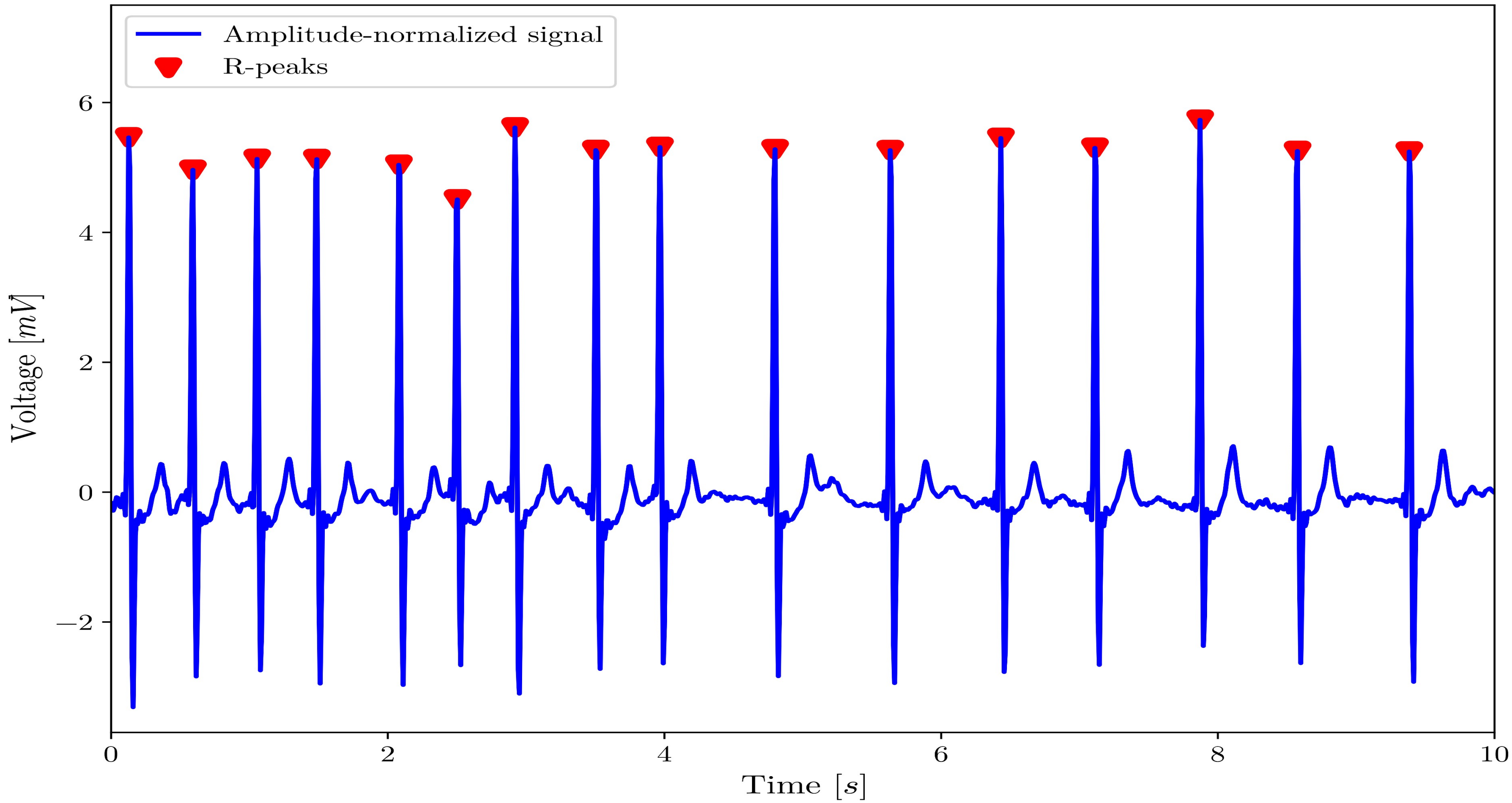


Fig. 8. Example of HRV for subject 221 (observation 9).

Heart rate variability (HRV)

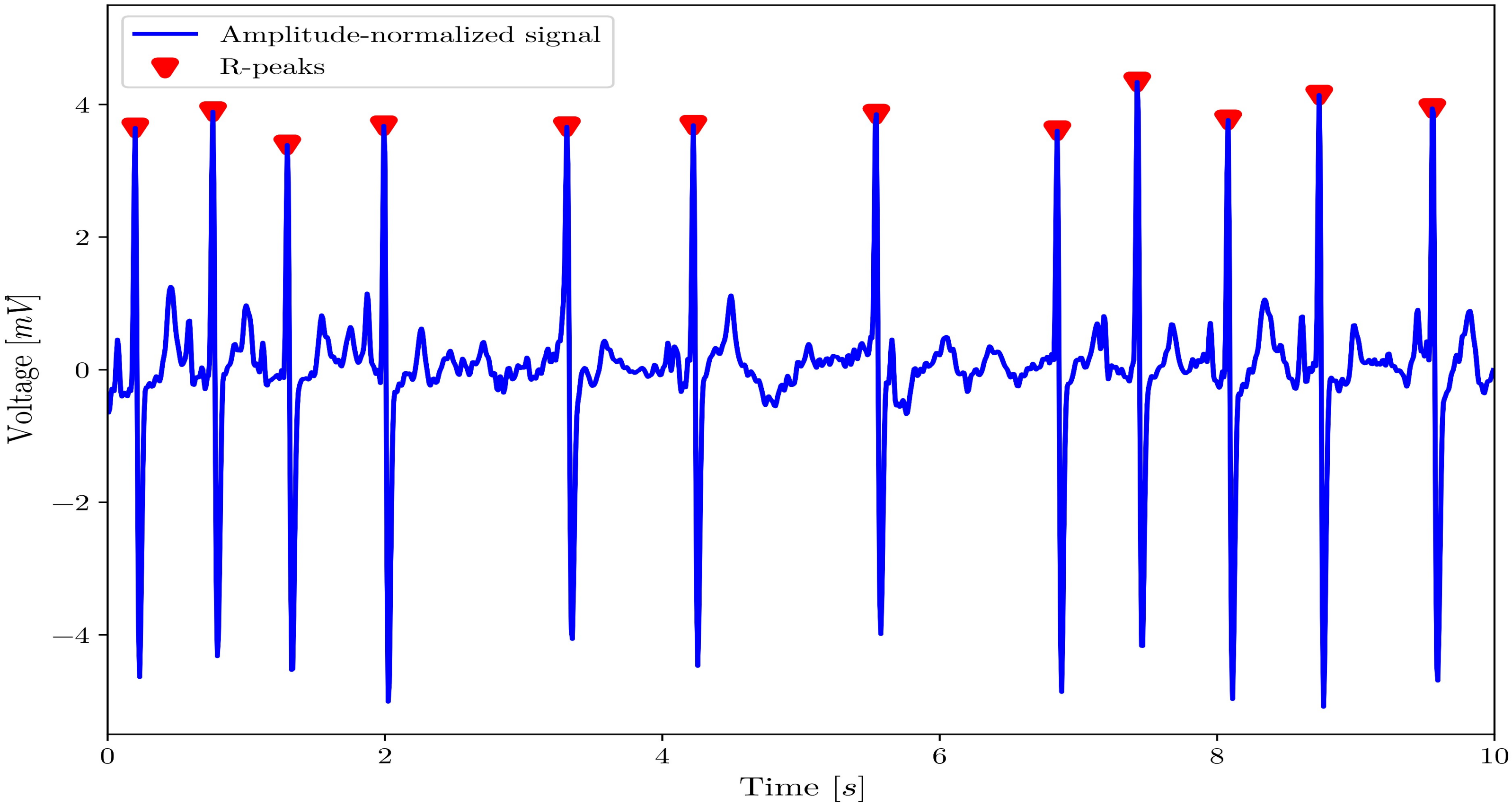


Fig. 9. Example of HRV for subject 394 (observation 7).

Heart rate variability (HRV)

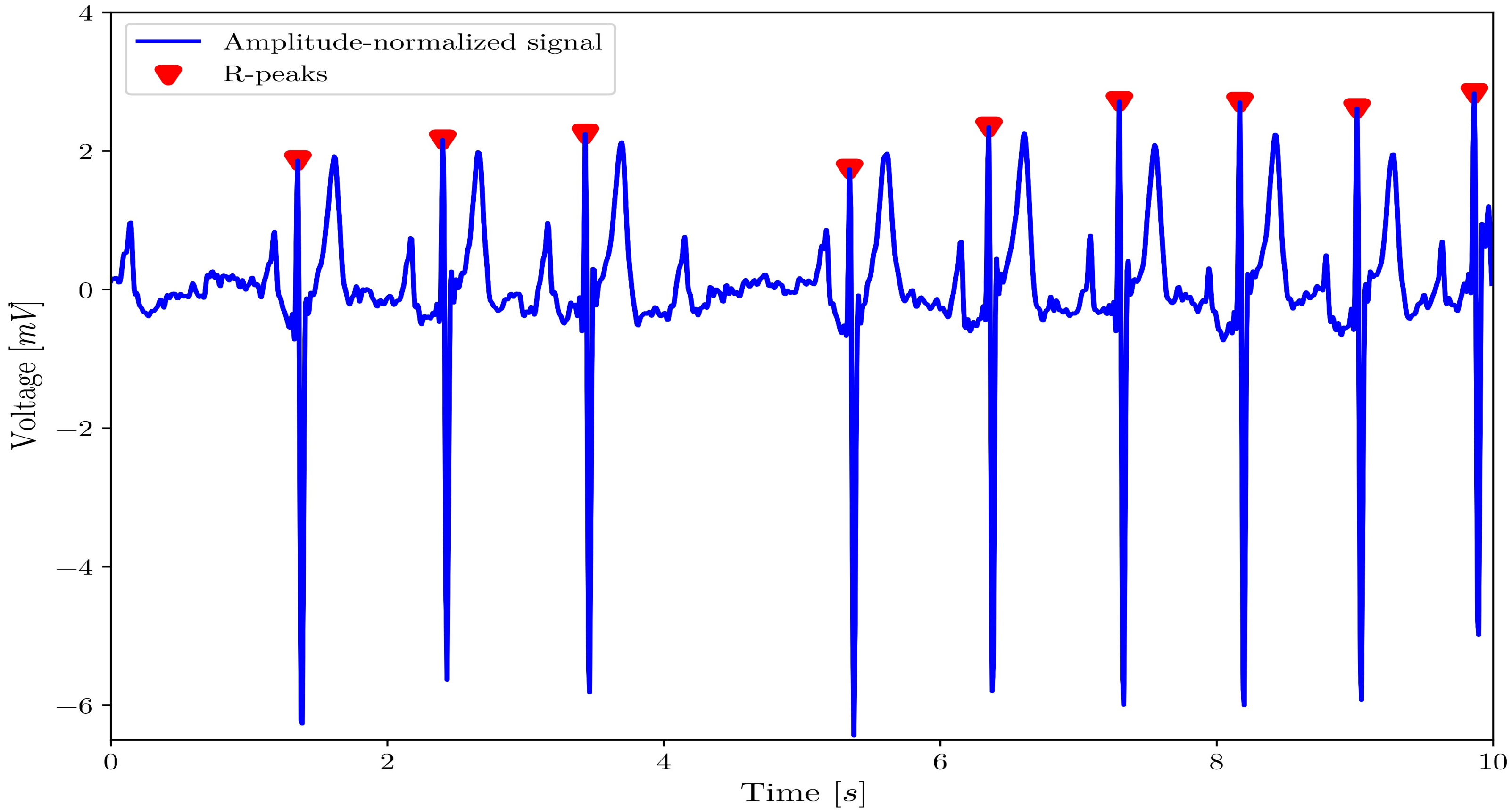


Fig. 10. Example of HRV for subject 207 (observation 10).

Scenarios of ECG signal variability

Physical Activities:

- Short-term variability, increasing heart rate.
- T wave approaches QRS complex.
- Daily activities like walking, running, stairs impact ECG signals.

Mental Stress:

- Short-term variability, can increase or decrease heart rate.
- Changes morphology of PQRST complex.
- Emotional (audiovisual stimuli) and cognitive (intellectual tasks) activities affect HRV.

Scenarios of ECG signal variability

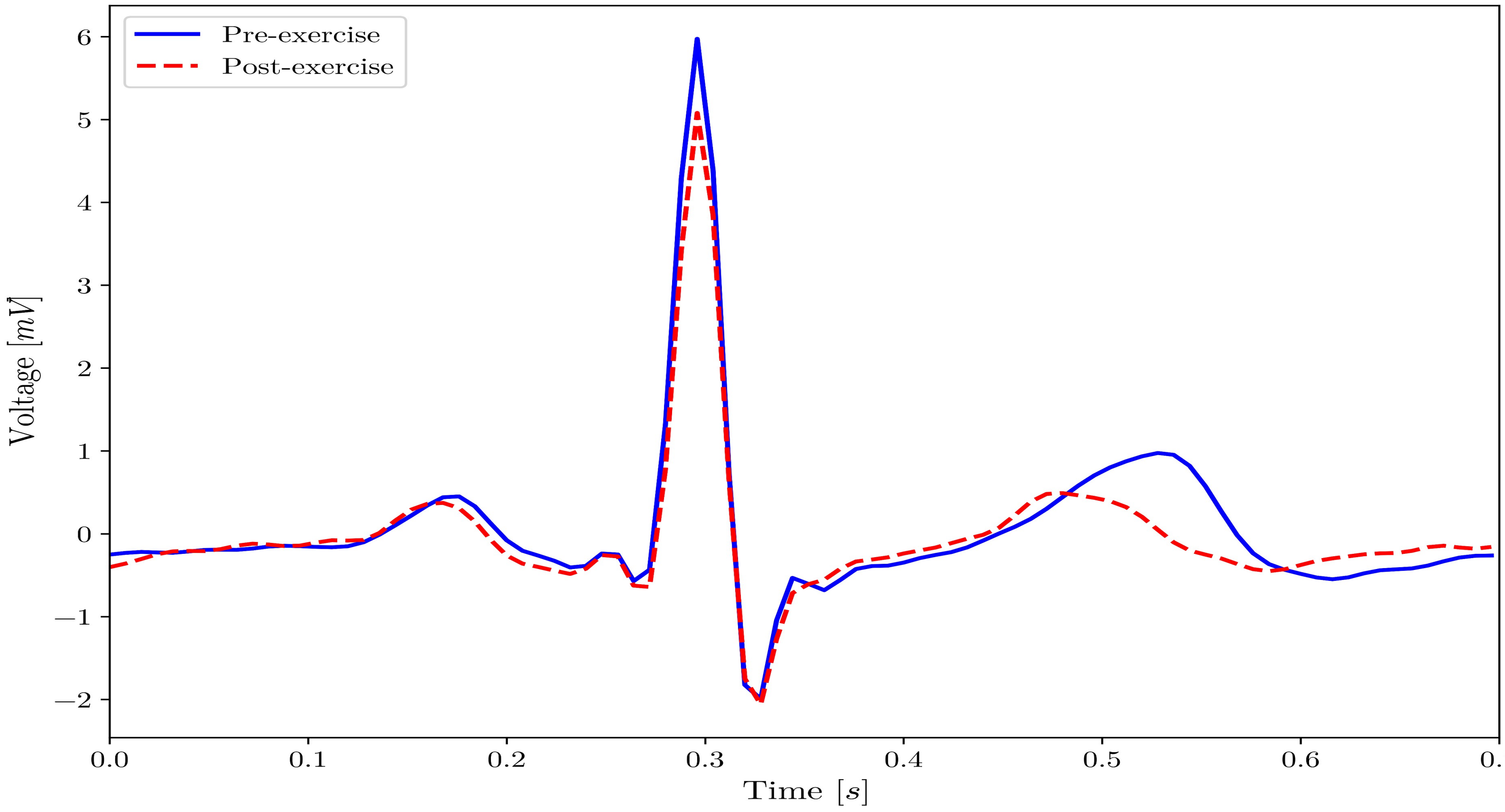


Fig. 11. Example of intra-user variability under physical activities for subject 550 (obs. 2 and 10).

Scenarios of ECG signal variability

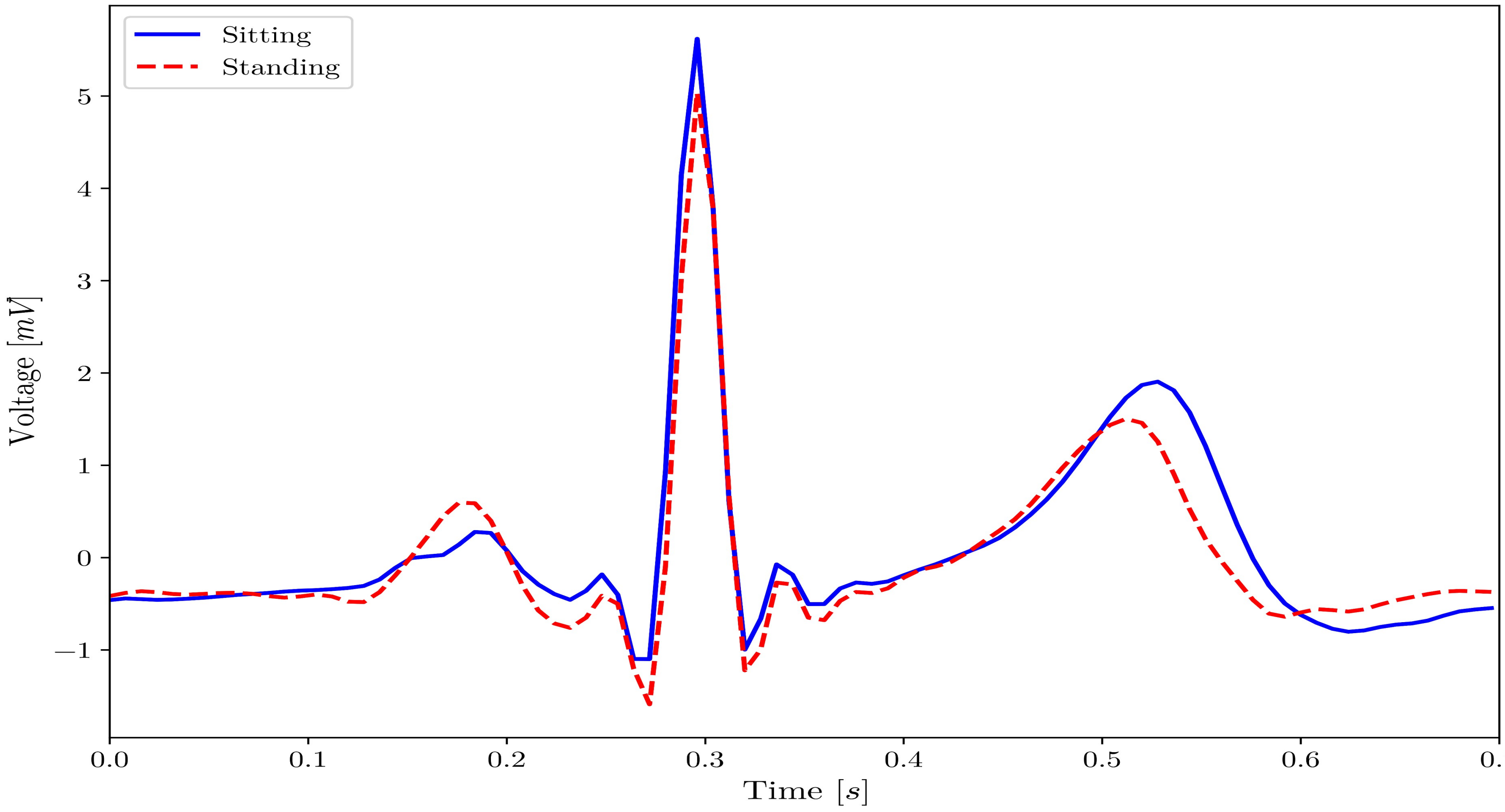


Fig. 12. Example of intra-user variability under acquisition postures for subject 561 (obs. 1 and 9).

Scenarios of ECG signal variability

Changes Over Time:

- Long-term variability from lifestyle, aging, and daily activities.
- Gradual anatomical changes in the heart.
- Temporal scenarios: ECG signals from the same person at different times (e.g., months apart).

Scenarios of ECG signal variability

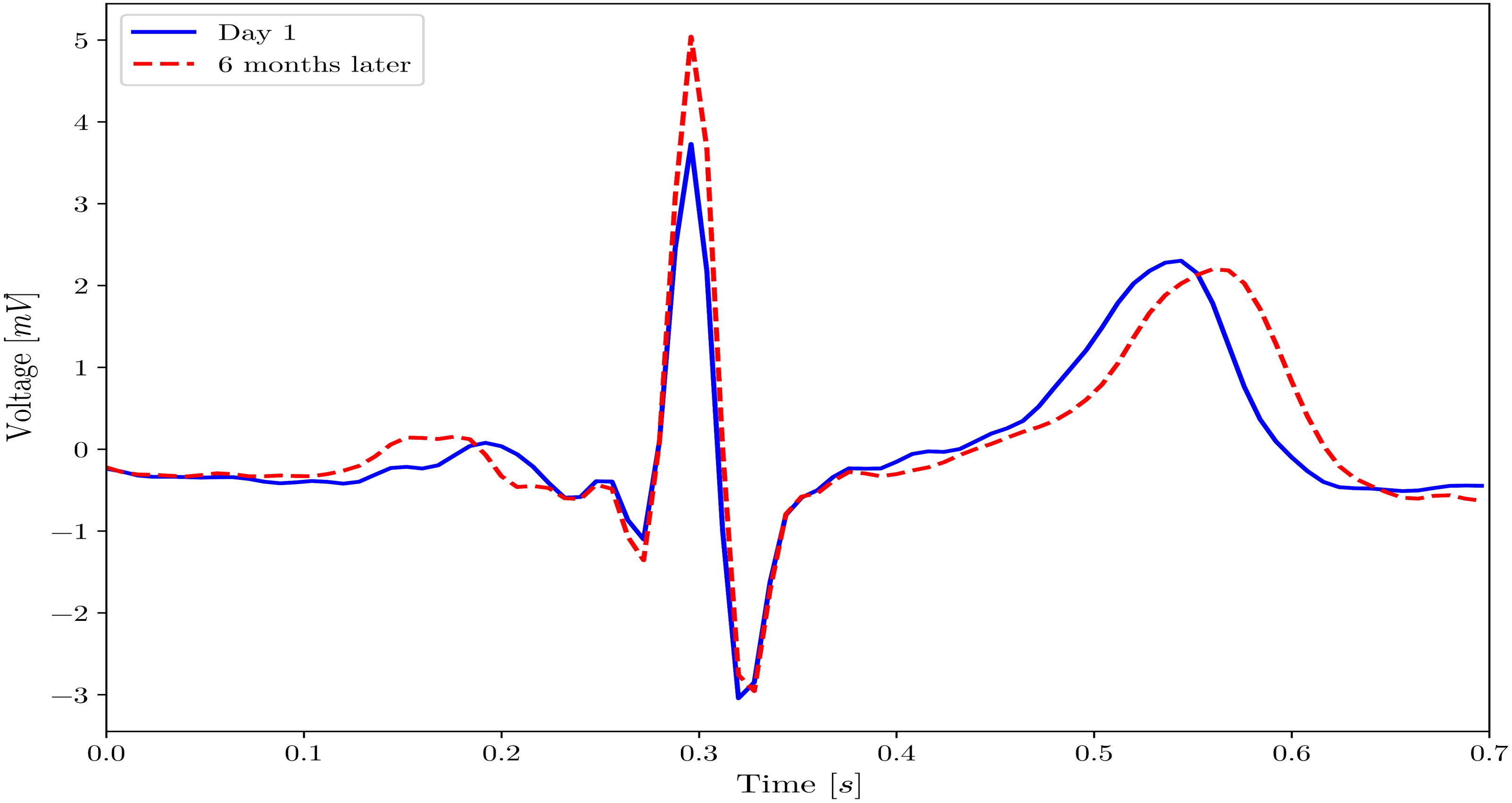


Fig. 13. Example of intra-user variability under changes over time for subject 2 (obs. 1 and 10).

Scenarios of ECG signal variability

Drug Effect:

- Long-term variability due to commercial drugs.
- Drugs can prolong QT interval, indicating proarrhythmic potential.
- Examples: ranolazine, verapamil, lopinavir, ritonavir, chloroquine, dofetilide, diltiazem.

Scenarios of ECG signal variability

Proarrhythmic effects of antiarrhythmic drugs

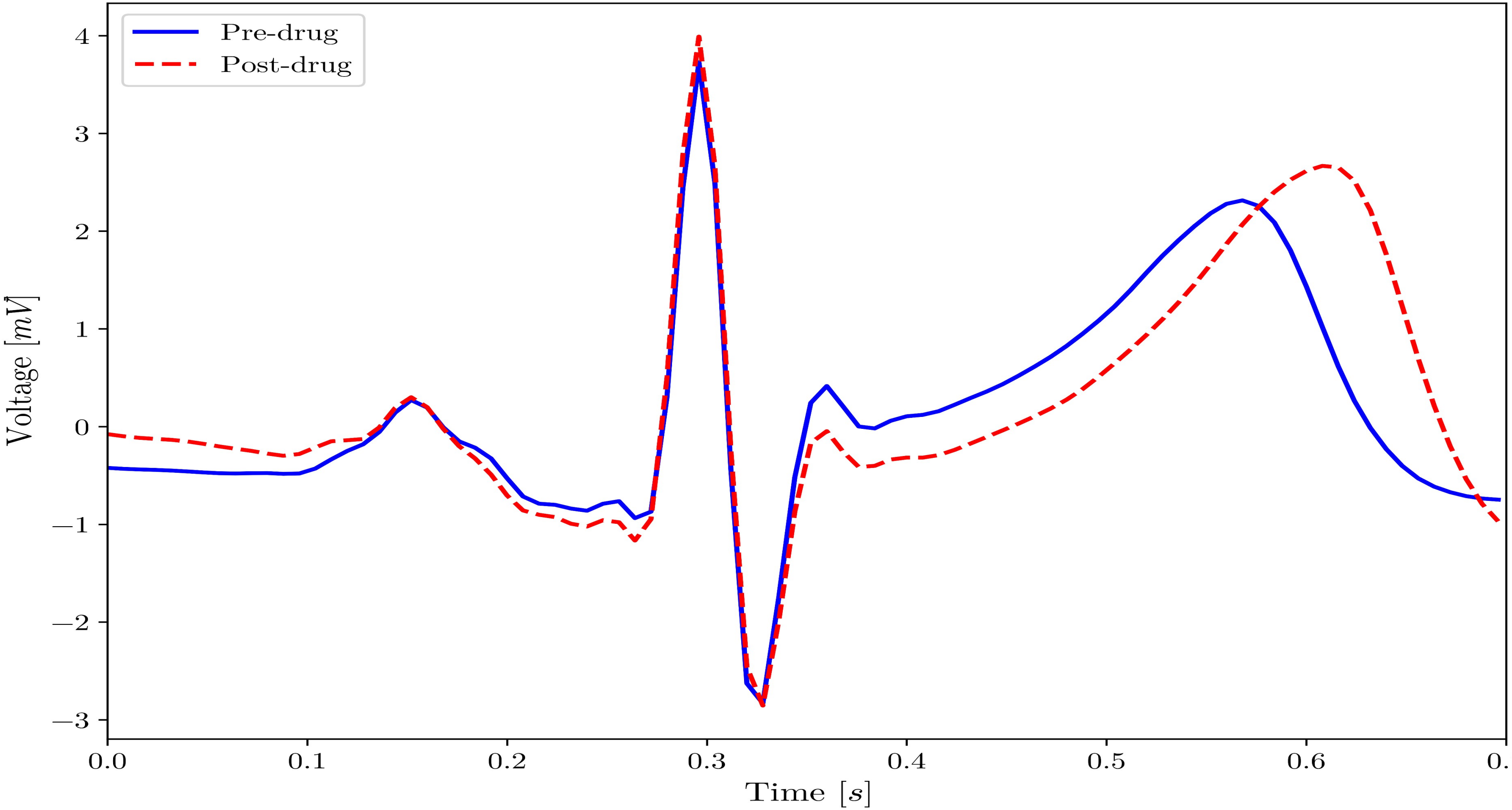


Fig. 14. Example of intra-user variability under drug effect (chloroquine) for subject 183 (obs. 1 and 10).

Scenarios of ECG signal variability

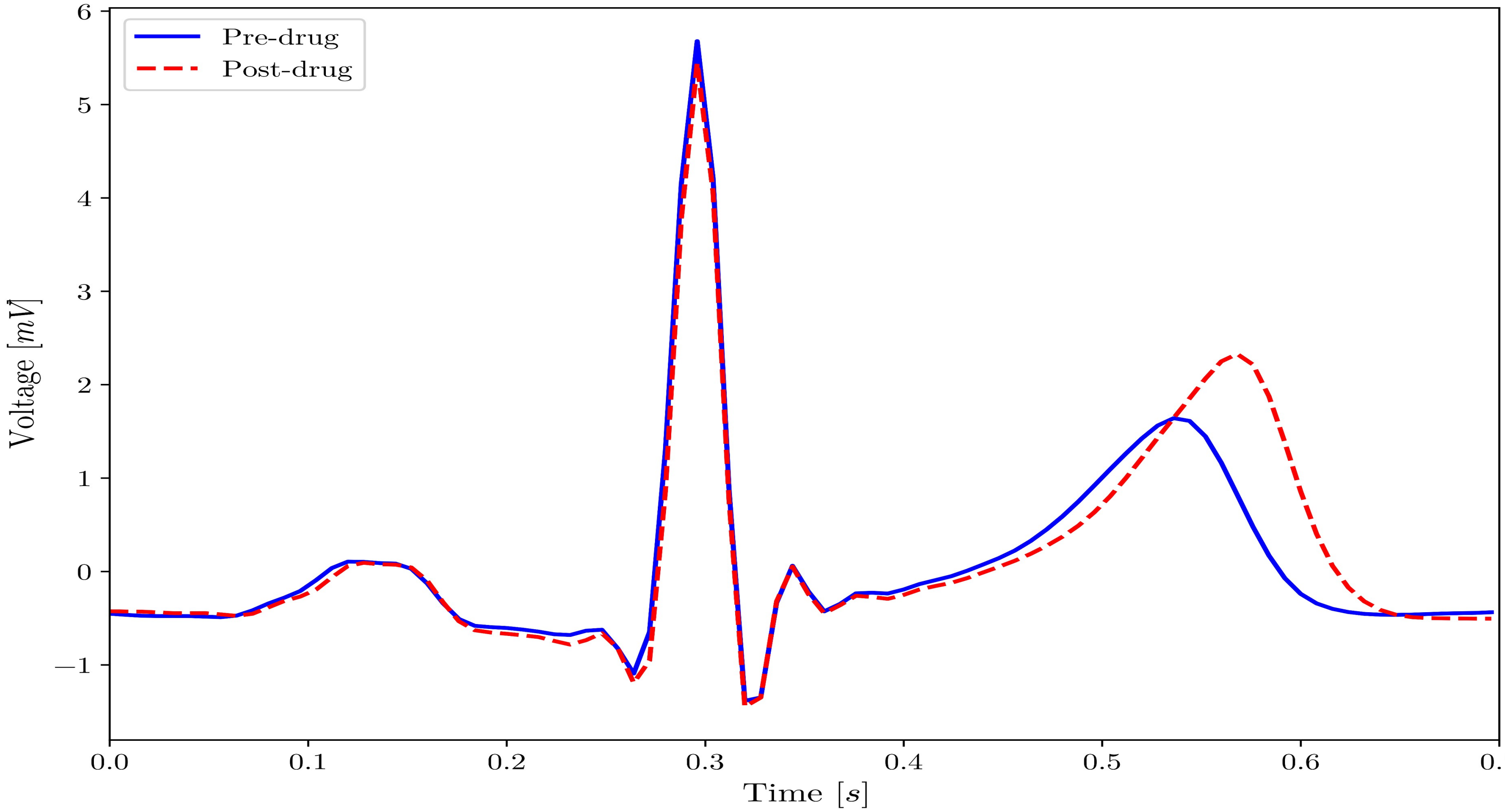


Fig. 15. Example of intra-user variability under drug effect (verapamil) for subject 196 (obs. 2 and 10).

Scenarios of ECG signal variability

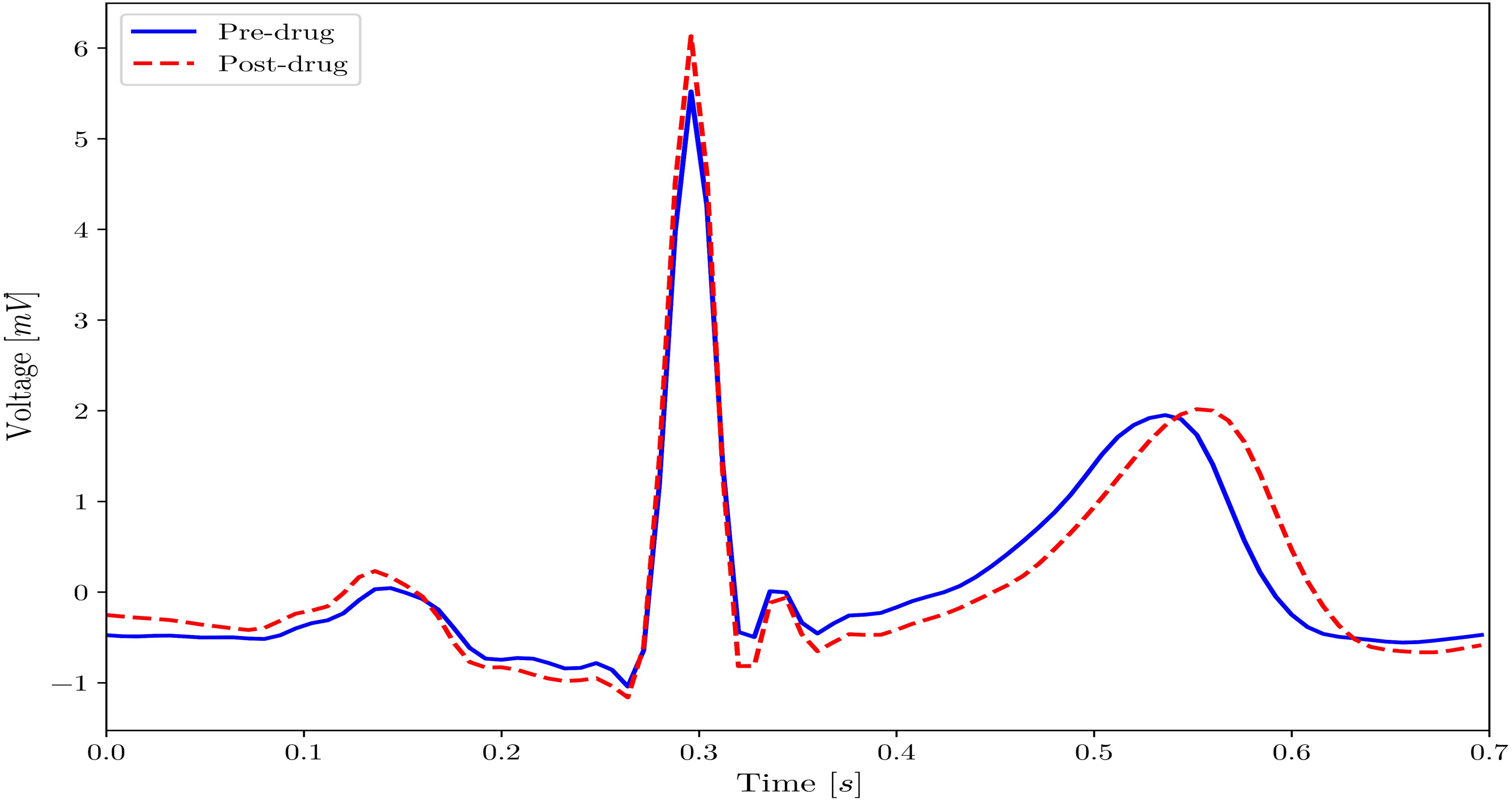


Fig. 16. Example of intra-user variability under drug effect (ranolazine) for subject 164 (obs. 2 and 10).

Scenarios of ECG signal variability

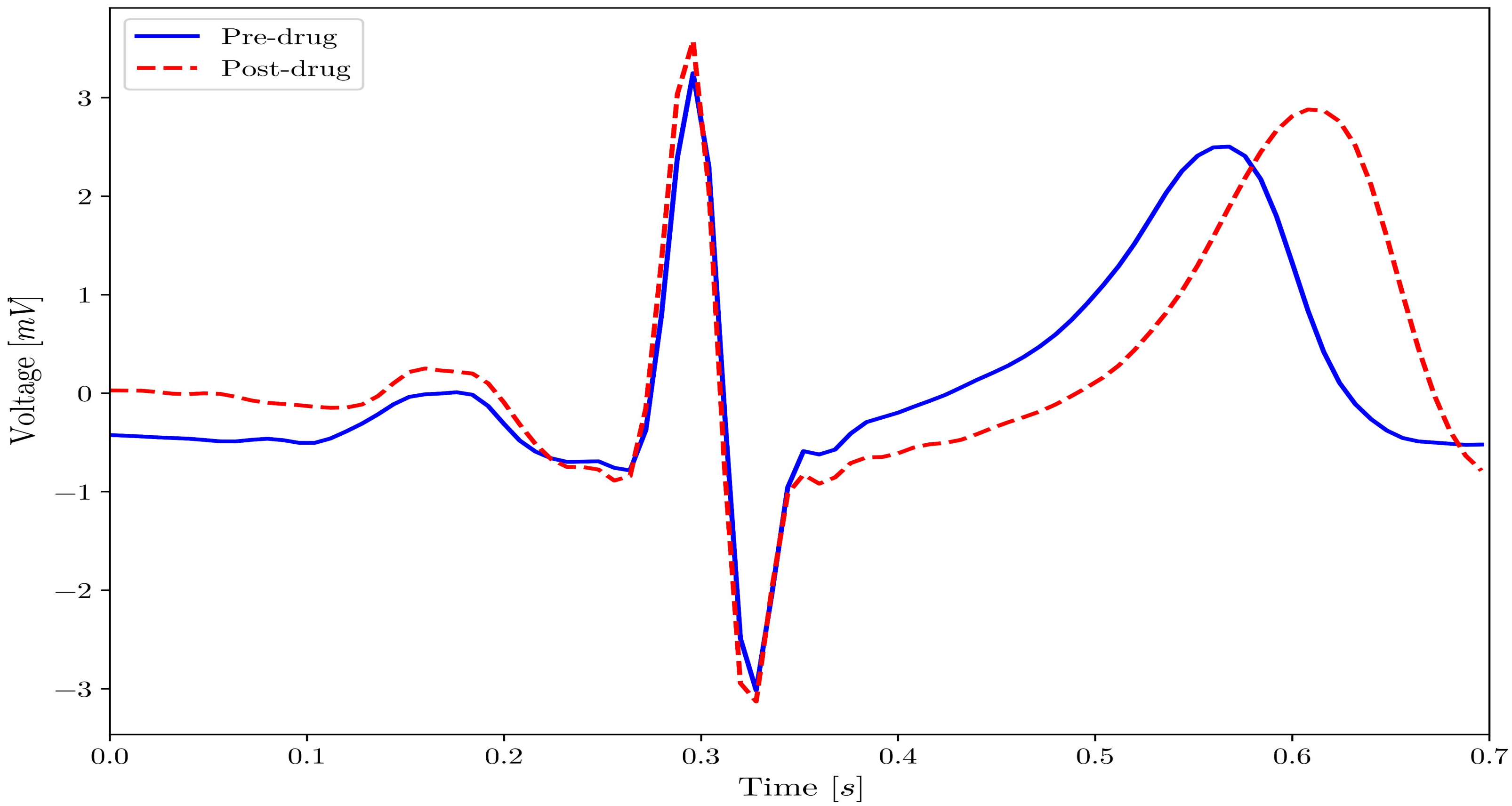


Fig. 17. Example of intra-user variability under drug effect (lopinavir+ritonavir) for subject 158 (obs. 2 and 5).

Scenarios of ECG signal variability

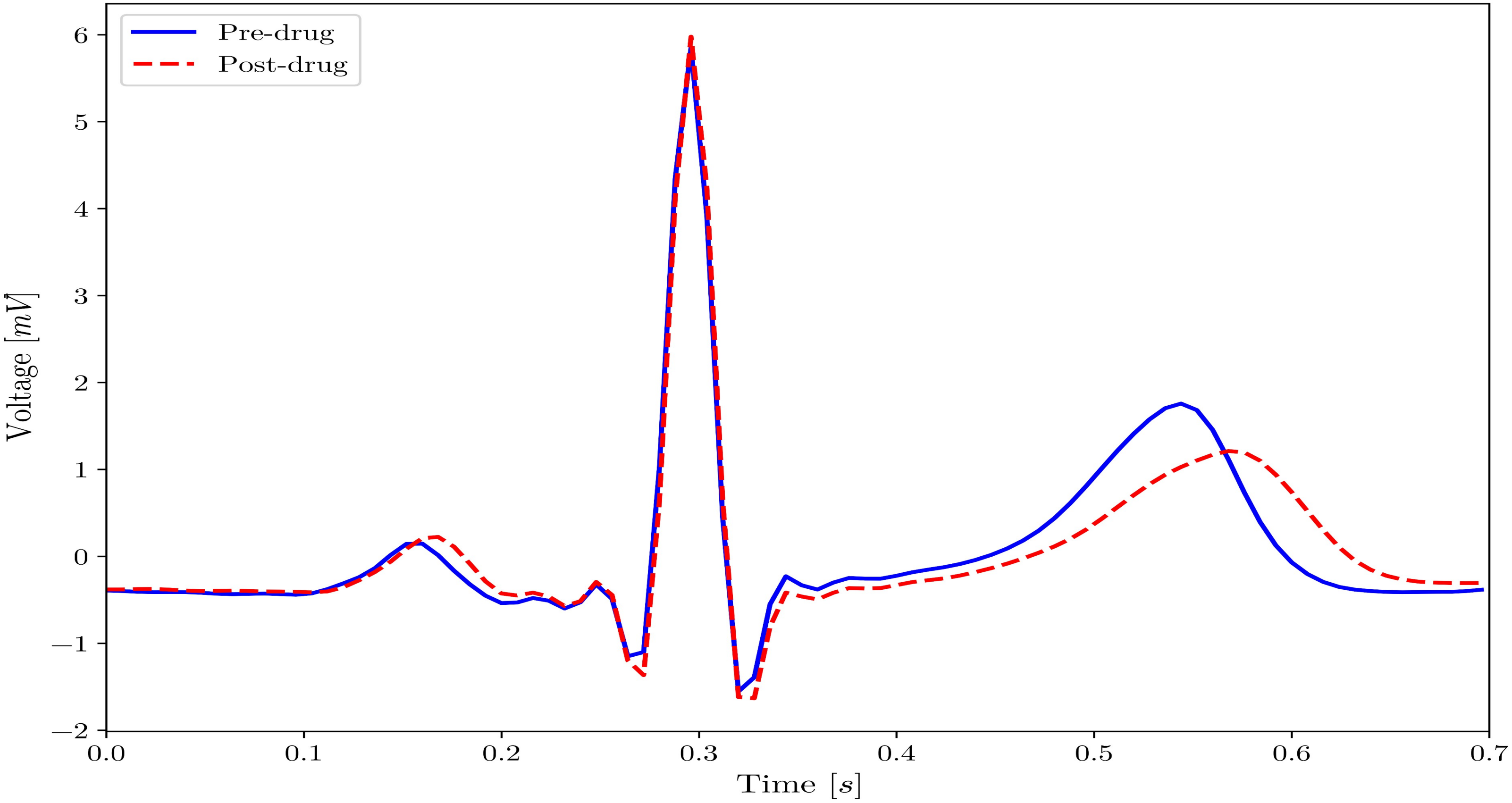


Fig. 18. Example of intra-user variability under drug effect (diltiazem+dofetilide) for subject 209 (obs. 2 and 10).

Scenarios of ECG signal variability

Effect of Heart Medical Antecedents:

- Long-term effects from past cardiac episodes, treatments, or procedures.
- Historical cardiac treatments or diseases alter ECG responses.

Scenarios of ECG signal variability

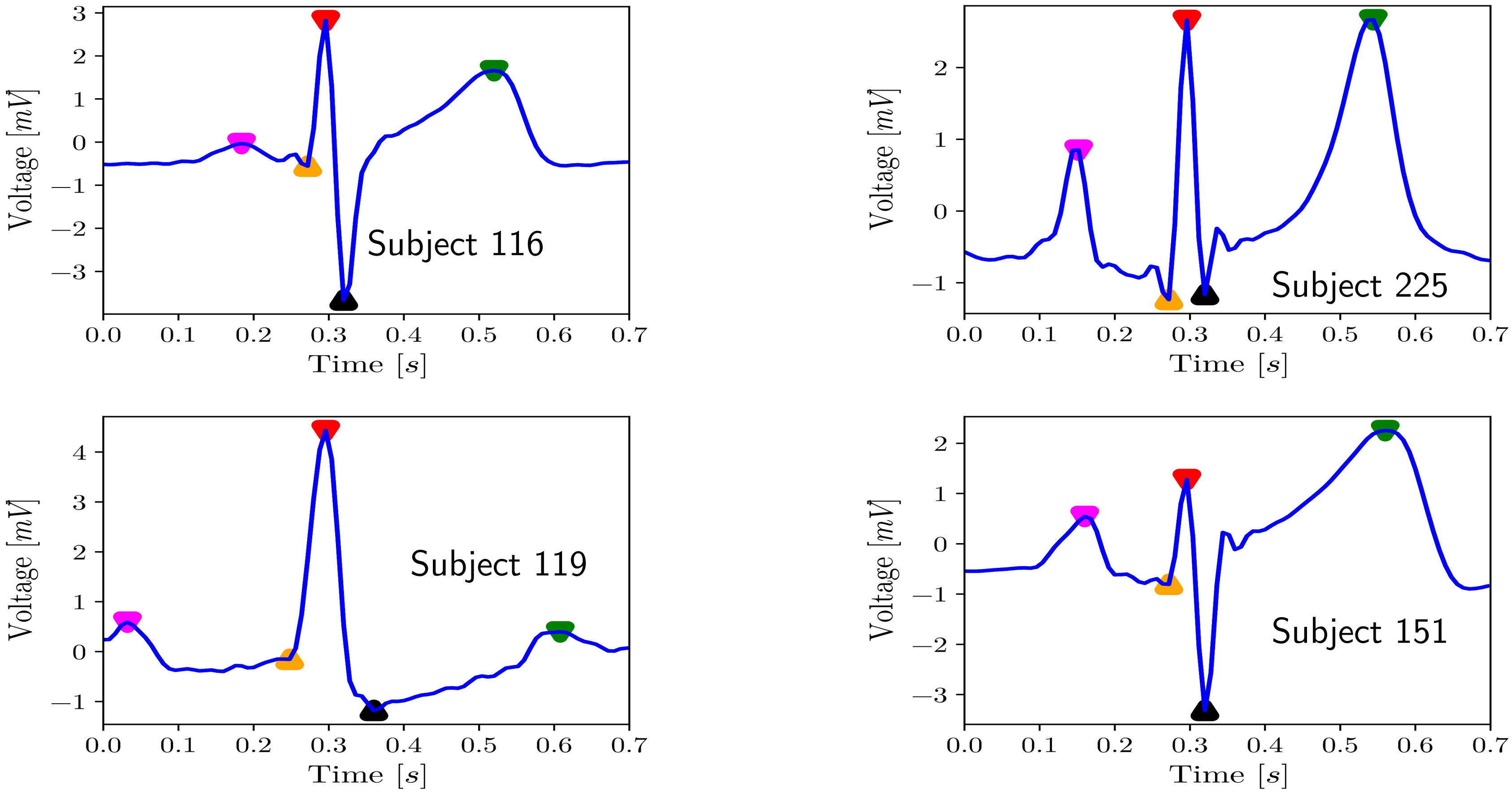


Fig. 19 [Continue]. Cardiac complex morphology (PQRST) for different subjects.

Scenarios of ECG signal variability

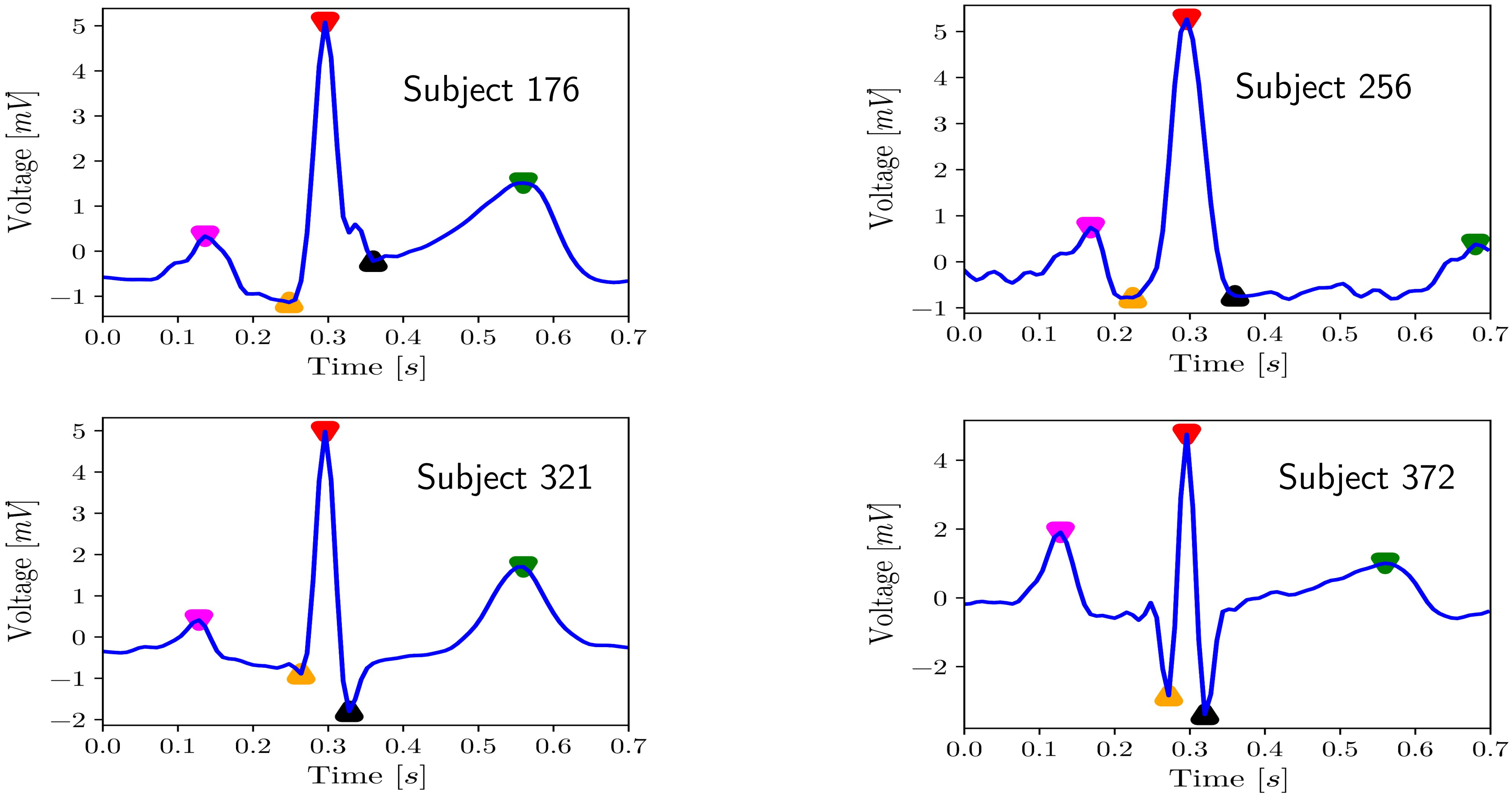


Fig. 19 [Continue]. Cardiac complex morphology (PQRST) for different subjects.

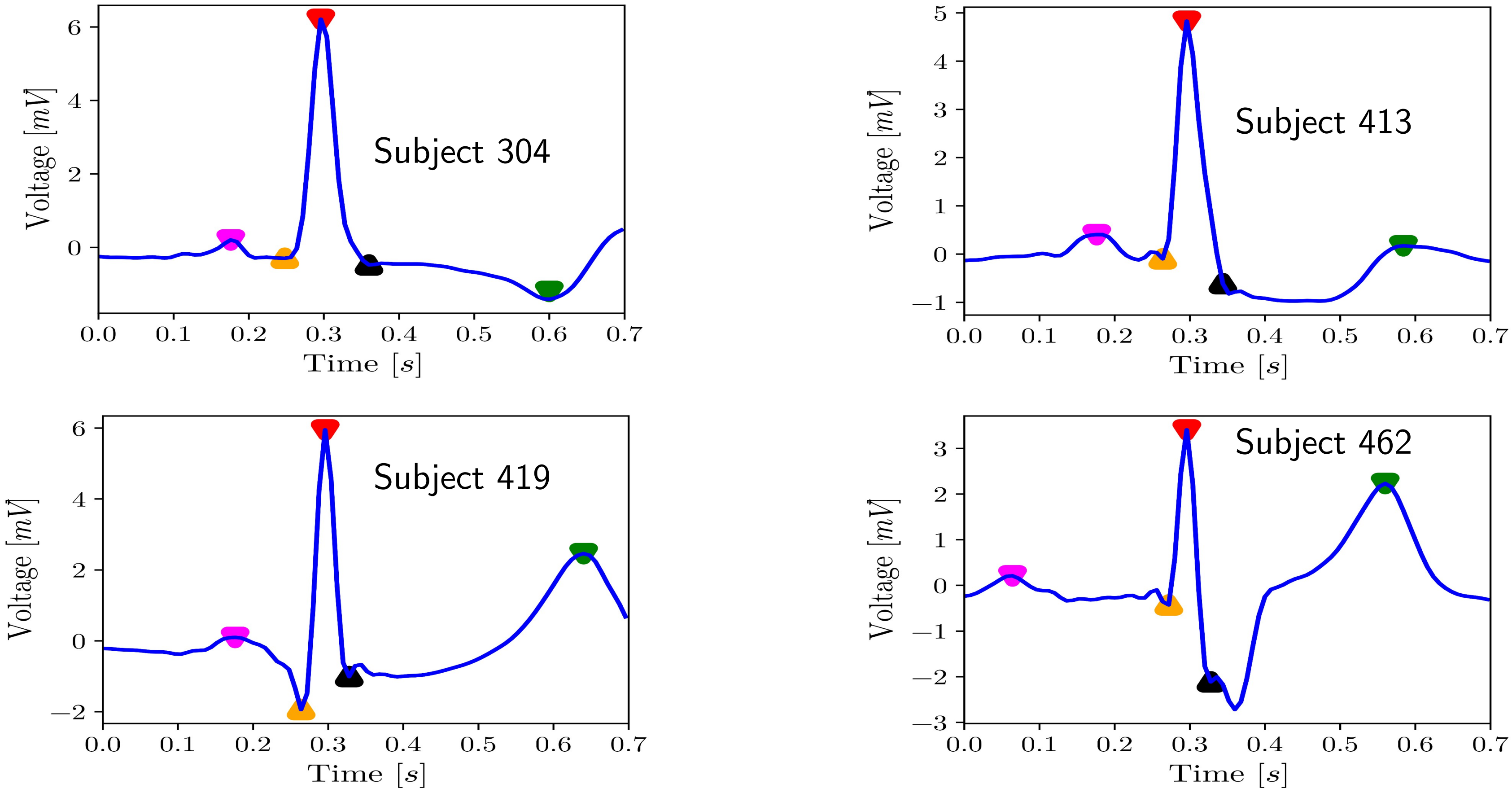


Fig. 19 [Continue]. Cardiac complex morphology (PQRST) for different subjects.

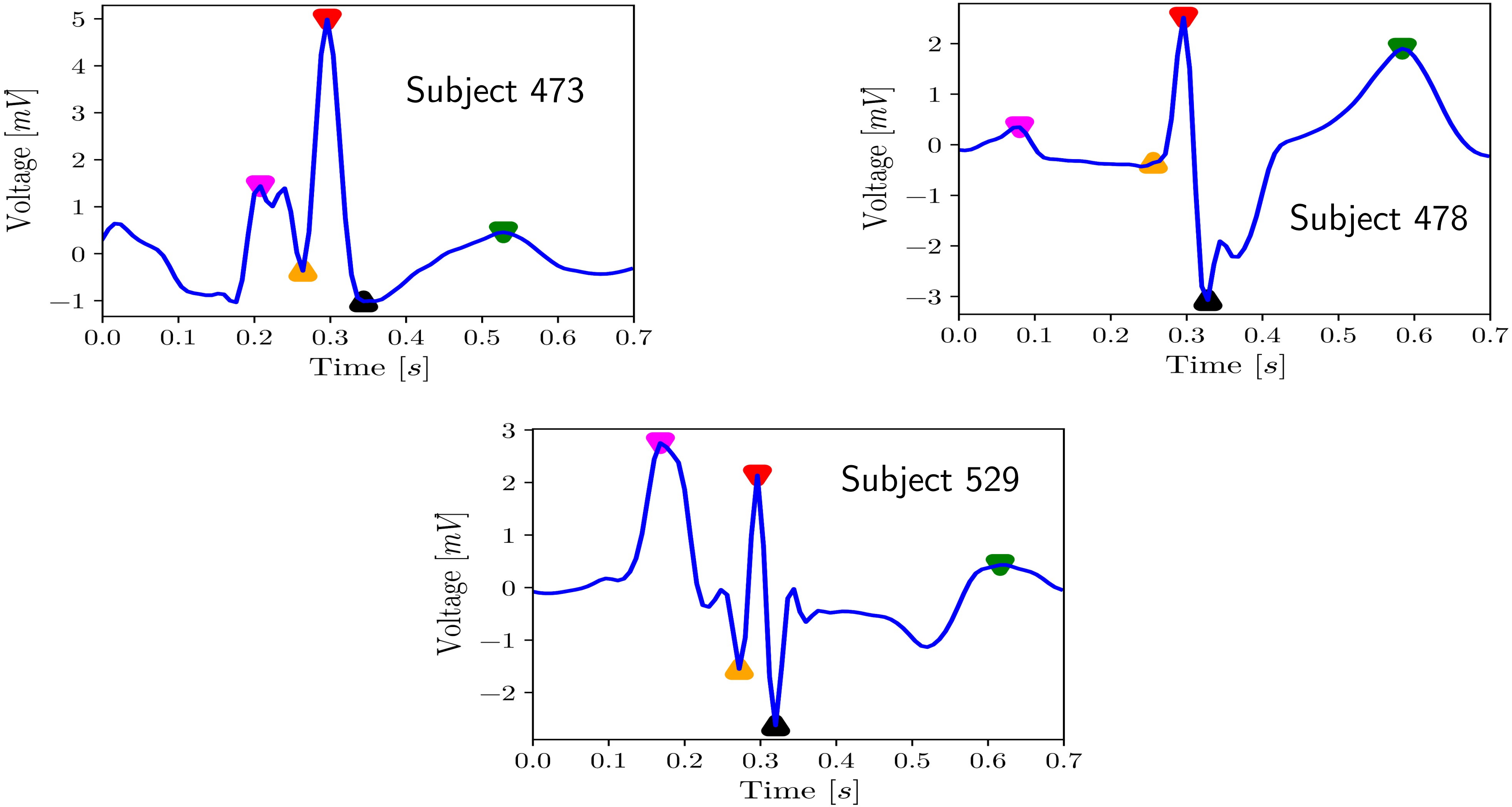


Fig. 19. Cardiac complex morphology (PQRST) for different subjects.

Consolidated lead-I ECG database

Table 1. Creation of the consolidated database

Item	Database	Subjects	Posture	Medical Antecedents	Variability scenario
1	owndb [1]	25	Sitting	Healthy	Physical Emotional
2	ecgiddb [2]	20	Sitting	Healthy	Time Physical Emotional
3	CYBHi [3]	91	Sitting	Healthy	Time Emotional
4	cebsdb [4]	20	Supine	Healthy	Time Emotional
5	ecgcipa [5]	60	Supine	Healthy	Time Drugs
6	staffiii [6]	102	Supine	Myocardial infarction	Time Medical procedures
7	ptbdb [7]	222	Supine	Healthy Cardiac diseases	Time Medical procedures
8	pttppgdb [8]	22	Supine Standing	Healthy	Time Physical

Consolidated lead-I ECG database

- The consolidated database contains 5620 recordings of ECG signals acquired in the upper extremities (lead-I) from 562 individuals (185 women and 377 men) with an age range of 17 to 100 years and an average age of 45 years.
- Each person has 10 lead-I ECG recordings of 10 seconds duration each.
- The database has a minimum, average, and maximum heart rate of 36 beats per minute (bpm), 71 bpm, and 162 bpm respectively.
- There are 289 healthy individuals, i.e., these individuals have no heart medical antecedents. Nonetheless, the database contains individuals with cardiac medical histories such as myocardial ischemia, myocardial infarction, dysrhythmia, cardiomyopathy, hypertrophy, bundle branch block, valvular heart disease, stable angina, and myocarditis.
- Pregnant women were not considered in this consolidated database, but there are 65 smokers.
- The consolidated database presents scenarios of temporal variability, physical, emotional, effects of drugs, and the effects of heart medical antecedents.
- The signals were acquired in different postures such as sitting, supine, and standing.
- All the signals in the consolidated database have a sampling frequency of 125 Hz and a resolution of 10 bits (VDD=3.3 V).

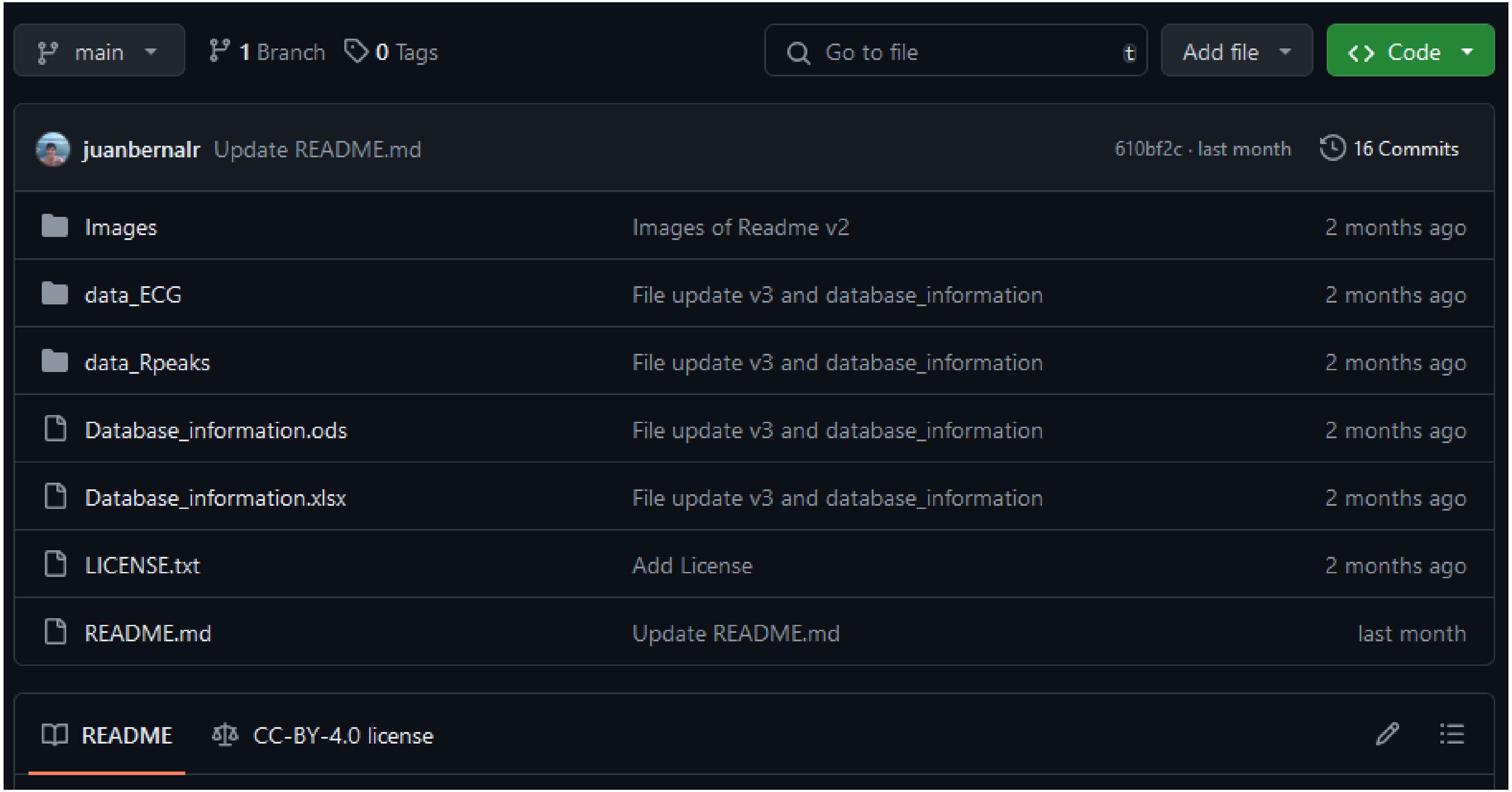
Consolidated lead-I ECG database

Database Composition:

- Variability Scenarios: Temporal, physical, emotional, drug effects, and cardiac medical antecedents.
- Acquisition Postures: Sitting, supine, standing.
- Technical Specifications: Sampling frequency: 125 Hz - Resolution: 10 bits.
- Detailed Metadata:
    - Sex, age, variability scenarios, acquisition posture.
    - Cardiac medical history, smoking habits.
    - Morphology-level observations of ECG signals.

- Public database on GitHub.

Consolidated lead-I ECG database



Conclusions

- An open-access database of lead-I ECG signals was created with practical variability scenarios (physical activity, mental stress, anatomical and physiological changes over time, and drug effects).
- The database has 562 individuals with 10 variability scenarios each. Each recording has 10 seconds.
- This database can be used for biometric systems or diagnostic systems.

References

[1] J. C. Bernal-Romero, J. M. Ruız-Echeverri, J. M. Ramírez-Cortes, P. Gomez-Gil, J. Rangel-Magdaleno, and I. Cruz-Vega, "On signal variability of ECG-based biometric system under practical considerations," in 2021 IEEE Mexican Humanitarian Technology Conference (MHTC), 2021, pp. 19–24.
[2] T. S. Lugovaya, "Biometric human identification based on electrocardiogram," Master's thesis, Faculty of Computing Technologies and Informatics, Saint Petersburg Electrotechnical University, Russian Federation, 2005.
[3] H. P. Da Silva, A. Lourenc¸o, A. Fred, N. Raposo, and M. Aires-de Sousa, "Check Your Biosignals Here: A new dataset for off-the-person ECG biometrics," Computer Methods and Programs in Biomedicine, vol. 113, no. 2, pp. 503–514, 2014.
[4] M. A. García-González, A. Argelagós-Palau, M. Fern´andez-Chimeno, and J. Ramos-Castro, "A comparison of heartbeat detectors for the seismocardiogram," in Computing in Cardiology 2013, 2013, pp. 461–464.
[5] J. Vicente, R. Zusterzeel, L. Johannesen, R. Ochoa-Jimenez, J. W. Mason, C. Sanabria, S. Kemp, P. T. Sager, V. Patel, M. K. Matta, J. Liu, J. Florian, C. Garnett, N. Stockbridge, and D. G. Strauss, "Assessment of multi-ion channel block in a phase I randomized study design:Results of the CiPA phase I ECG biomarker validation study," Clinical Pharmacology & Therapeutics, vol. 105, no. 4, pp. 943–953, 2019.

References

[6] J. P. Martınez, O. Pahlm, M. Ringborn, S. Warren, P. Laguna, and L. S¨ornmo, “The STAFF III database: ECGs recorded during acutely induced myocardial ischemia,” in 2017 Computing in Cardiology (CinC), 2017, pp. 1–4.
[7] R. Bousseljot, D. Kreiseler, and A. Schnabel, “Nutzung der EKG signaldatenbank cardiodat der PTB ¨uber das internet,” Biomedical Engineering / Biomedizinische Technik, vol. 40, no. s1, pp. 317–318, 1995.
[8] P. Mehrgardt, M. Khushi, S. Poon, and A. Withana, “Deep learning fused wearable pressure and PPG data for accurate heart rate monitoring,” IEEE Sensors Journal, vol. 21, no. 23, pp. 27 106–27 115, 2021.